

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
3 January 2002 (03.01.2002)

PCT

(10) International Publication Number  
WO 02/00174 A2

- (51) International Patent Classification<sup>7</sup>: **A61K**
- (21) International Application Number: PCT/US01/21065
- (22) International Filing Date: 28 June 2001 (28.06.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
- |            |                                |    |
|------------|--------------------------------|----|
| 09/606,421 | 28 June 2000 (28.06.2000)      | US |
| 09/630,940 | 2 August 2000 (02.08.2000)     | US |
| 09/643,597 | 21 August 2000 (21.08.2000)    | US |
| 09/662,786 | 15 September 2000 (15.09.2000) | US |
| 09/685,696 | 9 October 2000 (09.10.2000)    | US |
| 09/735,705 | 12 December 2000 (12.12.2000)  | US |
| 09/850,716 | 7 May 2001 (07.05.2001)        | US |

(71) Applicant (for all designated States except US): **CORIXA CORPORATION** [US/US]; 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **WANG, Tongtong** [US/US]; 8049 N.E. 28th Street, Medina, WA 98039 (US). **WANG, Aijun** [CN/US]; 3106 213th Place S.E., Issaquah, WA 98029 (US). **SKEIKY, Yasir, A., W.** [LB/US]; 15106 S.E. 47th Place, Bellevue, WA 98006 (US). **LI, Samuel, X.** [US/US]; 3608 175th Court N.E., Redmond, WA 98052 (US). **KALOS, Michael, D.** [US/US]; 8116 Dayton Ave. N., Seattle, WA 98103 (US). **HENDERSON, Robert, A.** [US/US]; 8904 192nd Street S.W., Edmonds, WA 98026 (US). **MCNEILL, Patricia, D.** [US/US]; 1333 South 290th Place, Federal Way, WA 98003 (US). **FANGER, Neil** [US/US]; 3648 Whitman Avenue N., A100, Seattle, WA 98103 (US). **RETTET, Marc, W.** [US/US]; 33402

N.E. 43rd Place, Carnation, WA 98014 (US). **MARN-ERAKIS, Margarita** [US/US]; 3444 36th Avenue W., Seattle, WA 98199 (US). **FANGER, Gary, Richard** [US/US]; 15906 29th Drive S.E., Mill Creek, WA 98012 (US). **VEDVICK, Thomas, S.** [US/US]; 124 S. 300th Place, Federal Way, WA 98003 (US). **CARTER, Darrick** [US/US]; 321 Summit Ave. E., Seattle, WA 98102 (US). **WATANABE, Yoshihiro** [JP/US]; 2266 78th Avenue S.E., Mercer Island, WA 98040 (US). **PECKHAM, David, W.** [US/US]; 903 9th Avenue, Apt. #31, Seattle, WA 98104 (US).

(74) Agents: **VERNA, James, M.**; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 et al. (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MD, SZ, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, particularly lung cancer, are disclosed. Illustrative compositions comprise one or more lung tumor polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly lung cancer.



WO 02/00174 A2

## COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER

### TECHNICAL FIELD OF THE INVENTION

5           The present invention relates generally to therapy and diagnosis of cancer, such as lung cancer. The invention is more specifically related to polypeptides, comprising at least a portion of a lung tumor protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides are useful in pharmaceutical compositions, *e.g.*, vaccines, and other compositions for the diagnosis and treatment of  
10 lung cancer.

### BACKGROUND OF THE INVENTION

Cancer is a significant health problem throughout the world. Although advances have been made in detection and therapy of cancer, no vaccine or other universally successful method for prevention and/or treatment is currently available.  
15 Current therapies, which are generally based on a combination of chemotherapy or surgery and radiation, continue to prove inadequate in many patients.

Lung cancer is the primary cause of cancer death among both men and women in the U.S., with an estimated 172,000 new cases being reported in 1994. The five-year survival rate among all lung cancer patients, regardless of the stage of disease  
20 at diagnosis, is only 13%. This contrasts with a five-year survival rate of 46% among cases detected while the disease is still localized. However, only 16% of lung cancers are discovered before the disease has spread.

In spite of considerable research into therapies for these and other cancers, lung cancer remains difficult to diagnose and treat effectively. Accordingly,  
25 there is a need in the art for improved methods for detecting and treating such cancers. The present invention fulfills these needs and further provides other related advantages.

## SUMMARY OF THE INVENTION

In one aspect, the present invention provides polynucleotide compositions comprising a sequence selected from the group consisting of:

- (a) sequences provided in SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467;
- (b) complements of the sequences provided in SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467;
- (c) sequences consisting of at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 75 and 100 contiguous residues of a sequence provided in SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467;
- (d) sequences that hybridize to a sequence provided in SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467, under moderate or highly stringent conditions;

(e) sequences having at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identity to a sequence of SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467; and

(f) degenerate variants of a sequence provided in SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467.

In one preferred embodiment, the polynucleotide compositions of the invention are expressed in at least about 20%, more preferably in at least about 30%, and most preferably in at least about 50% of lung tumors samples tested, at a level that is at least about 2-fold, preferably at least about 5-fold, and most preferably at least about 10-fold higher than that for normal tissues.

The present invention, in another aspect, provides polypeptide compositions comprising an amino acid sequence that is encoded by a polynucleotide sequence described above.

The present invention further provides polypeptide compositions comprising an amino acid sequence selected from the group consisting of sequences recited in SEQ ID NOs:152, 155, 156, 165, 166, 169, 170, 172, 174, 176, 226-252, 338-344, 346, 350, 357, 361, 363, 365, 367, 369, 376-382, 387-419, 423, 427, 430, 433, 441, 443, 446, 449 and 451-466, .

In certain preferred embodiments, the polypeptides and/or polynucleotides of the present invention are immunogenic, *i.e.*, they are capable of eliciting an immune response, particularly a humoral and/or cellular immune response, as further described herein.



The present invention further provides fragments, variants and/or derivatives of the disclosed polypeptide and/or polynucleotide sequences, wherein the fragments, variants and/or derivatives preferably have a level of immunogenic activity of at least about 50%, preferably at least about 70% and more preferably at least about 90% of the level of immunogenic activity of a polypeptide sequence set forth in SEQ ID NOs:152, 155, 156, 165, 166, 169, 170, 172, 174, 176, 226-252, 338-344, 346, 350, 357, 361, 363, 365, 367, 369, 376-382, 387-419, 423, 427, 430, 433, 441, 443, 446, 449 and 451-466, or a polypeptide sequence encoded by a polynucleotide sequence set forth in SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467.

The present invention further provides polynucleotides that encode a polypeptide described above, expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, the pharmaceutical compositions, e.g., vaccine compositions, are provided for prophylactic or therapeutic applications. Such compositions generally comprise an immunogenic polypeptide or polynucleotide of the invention and an immunostimulant, such as an adjuvant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a polypeptide of the present invention, or a fragment thereof; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Illustrative

antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, pharmaceutical compositions are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above  
5 and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins, typically in the form of pharmaceutical compositions, *e.g.*, vaccine compositions, comprising a physiologically acceptable carrier and/or an  
10 immunostimulant. The fusions proteins may comprise multiple immunogenic polypeptides or portions/variants thereof, as described herein, and may further comprise one or more polypeptide segments for facilitating the expression, purification and/or immunogenicity of the polypeptide(s).

Within further aspects, the present invention provides methods for  
15 stimulating an immune response in a patient, preferably a T cell response in a human patient, comprising administering a pharmaceutical composition described herein. The patient may be afflicted with lung cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

Within further aspects, the present invention provides methods for  
20 inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition as recited above. The patient may be afflicted with lung cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

The present invention further provides, within other aspects, methods for  
25 removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a polypeptide of the present invention, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating  
5 and/or expanding T cells specific for a polypeptide of the present invention, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells  
10 prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the  
15 development of a cancer in a patient, comprising the steps of: (a) incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of polypeptide disclosed herein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount  
20 of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for determining the presence or absence of a cancer, preferably a lung cancer, in a patient  
25 comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent  
30 is an antibody, more preferably a monoclonal antibody.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

10 The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a polypeptide of the present invention; (b) detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a polypeptide of the present invention; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount

detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as  
5 diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description and attached drawings. All references disclosed herein are hereby incorporated by reference in their entirety as if  
10 each was incorporated individually.

#### SEQUENCE IDENTIFIERS

- SEQ ID NO:1 is the determined cDNA sequence for LST-S1-2  
SEQ ID NO:2 is the determined cDNA sequence for LST-S1-28  
SEQ ID NO:3 is the determined cDNA sequence for LST-S1-90  
15 SEQ ID NO:4 is the determined cDNA sequence for LST-S1-144  
SEQ ID NO:5 is the determined cDNA sequence for LST-S1-133  
SEQ ID NO:6 is the determined cDNA sequence for LST-S1-169  
SEQ ID NO:7 is the determined cDNA sequence for LST-S2-6  
SEQ ID NO:8 is the determined cDNA sequence for LST-S2-11  
20 SEQ ID NO:9 is the determined cDNA sequence for LST-S2-17  
SEQ ID NO:10 is the determined cDNA sequence for LST-S2-25  
SEQ ID NO:11 is the determined cDNA sequence for LST-S2-39  
SEQ ID NO:12 is a first determined cDNA sequence for LST-S2-43  
SEQ ID NO:13 is a second determined cDNA sequence for LST-S2-43  
25 SEQ ID NO:14 is the determined cDNA sequence for LST-S2-65  
SEQ ID NO:15 is the determined cDNA sequence for LST-S2-68  
SEQ ID NO:16 is the determined cDNA sequence for LST-S2-72  
SEQ ID NO:17 is the determined cDNA sequence for LST-S2-74  
SEQ ID NO:18 is the determined cDNA sequence for LST-S2-103

- SEQ ID NO:19 is the determined cDNA sequence for LST-S2-N1-1F  
SEQ ID NO:20 is the determined cDNA sequence for LST-S2-N1-2A  
SEQ ID NO:21 is the determined cDNA sequence for LST-S2-N1-4H  
SEQ ID NO:22 is the determined cDNA sequence for LST-S2-N1-5A  
5 SEQ ID NO:23 is the determined cDNA sequence for LST-S2-N1-6B  
SEQ ID NO:24 is the determined cDNA sequence for LST-S2-N1-7B  
SEQ ID NO:25 is the determined cDNA sequence for LST-S2-N1-7H  
SEQ ID NO:26 is the determined cDNA sequence for LST-S2-N1-8A  
SEQ ID NO:27 is the determined cDNA sequence for LST-S2-N1-8D  
10 SEQ ID NO:28 is the determined cDNA sequence for LST-S2-N1-9A  
SEQ ID NO:29 is the determined cDNA sequence for LST-S2-N1-9E  
SEQ ID NO:30 is the determined cDNA sequence for LST-S2-N1-10A  
SEQ ID NO:31 is the determined cDNA sequence for LST-S2-N1-10G  
SEQ ID NO:32 is the determined cDNA sequence for LST-S2-N1-11A  
15 SEQ ID NO:33 is the determined cDNA sequence for LST-S2-N1-12C  
SEQ ID NO:34 is the determined cDNA sequence for LST-S2-N1-12E  
SEQ ID NO:35 is the determined cDNA sequence for LST-S2-B1-3D  
SEQ ID NO:36 is the determined cDNA sequence for LST-S2-B1-6C  
SEQ ID NO:37 is the determined cDNA sequence for LST-S2-B1-5D  
20 SEQ ID NO:38 is the determined cDNA sequence for LST-S2-B1-5F  
SEQ ID NO:39 is the determined cDNA sequence for LST-S2-B1-6G  
SEQ ID NO:40 is the determined cDNA sequence for LST-S2-B1-8A  
SEQ ID NO:41 is the determined cDNA sequence for LST-S2-B1-8D  
SEQ ID NO:42 is the determined cDNA sequence for LST-S2-B1-10A  
25 SEQ ID NO:43 is the determined cDNA sequence for LST-S2-B1-9B  
SEQ ID NO:44 is the determined cDNA sequence for LST-S2-B1-9F  
SEQ ID NO:45 is the determined cDNA sequence for LST-S2-B1-12D  
SEQ ID NO:46 is the determined cDNA sequence for LST-S2-I2-2B  
SEQ ID NO:47 is the determined cDNA sequence for LST-S2-I2-5F  
30 SEQ ID NO:48 is the determined cDNA sequence for LST-S2-I2-6B

- SEQ ID NO:49 is the determined cDNA sequence for LST-S2-I2-7F  
SEQ ID NO:50 is the determined cDNA sequence for LST-S2-I2-8G  
SEQ ID NO:51 is the determined cDNA sequence for LST-S2-I2-9E  
SEQ ID NO:52 is the determined cDNA sequence for LST-S2-I2-12B  
5 SEQ ID NO:53 is the determined cDNA sequence for LST-S2-H2-2C  
SEQ ID NO:54 is the determined cDNA sequence for LST-S2-H2-1G  
SEQ ID NO:55 is the determined cDNA sequence for LST-S2-H2-4G  
SEQ ID NO:56 is the determined cDNA sequence for LST-S2-H2-3H  
SEQ ID NO:57 is the determined cDNA sequence for LST-S2-H2-5G  
10 SEQ ID NO:58 is the determined cDNA sequence for LST-S2-H2-9B  
SEQ ID NO:59 is the determined cDNA sequence for LST-S2-H2-10H  
SEQ ID NO:60 is the determined cDNA sequence for LST-S2-H2-12D  
SEQ ID NO: 61 is the determined cDNA sequence for LST-S3-2  
SEQ ID NO: 62 is the determined cDNA sequence for LST-S3-4  
15 SEQ ID NO: 63 is the determined cDNA sequence for LST-S3-7  
SEQ ID NO: 64 is the determined cDNA sequence for LST-S3-8  
SEQ ID NO: 65 is the determined cDNA sequence for LST-S3-12  
SEQ ID NO: 66 is the determined cDNA sequence for LST-S3-13  
SEQ ID NO: 67 is the determined cDNA sequence for LST-S3-14  
20 SEQ ID NO: 68 is the determined cDNA sequence for LST-S3-16  
SEQ ID NO: 69 is the determined cDNA sequence for LST-S3-21  
SEQ ID NO: 70 is the determined cDNA sequence for LST-S3-22  
SEQ ID NO: 71 is the determined cDNA sequence for LST-S1-7  
SEQ ID NO: 72 is the determined cDNA sequence for LST-S1-A-1E  
25 SEQ ID NO: 73 is the determined cDNA sequence for LST-S1-A-1G  
SEQ ID NO: 74 is the determined cDNA sequence for LST-S1-A-3E  
SEQ ID NO: 75 is the determined cDNA sequence for LST-S1-A-4E  
SEQ ID NO: 76 is the determined cDNA sequence for LST-S1-A-6D  
SEQ ID NO: 77 is the determined cDNA sequence for LST-S1-A-8D  
30 SEQ ID NO: 78 is the determined cDNA sequence for LST-S1-A-10A

- SEQ ID NO: 79 is the determined cDNA sequence for LST-S1-A-10C  
SEQ ID NO: 80 is the determined cDNA sequence for LST-S1-A-9D  
SEQ ID NO: 81 is the determined cDNA sequence for LST-S1-A-10D  
SEQ ID NO: 82 is the determined cDNA sequence for LST-S1-A-9H  
5 SEQ ID NO: 83 is the determined cDNA sequence for LST-S1-A-11D  
SEQ ID NO: 84 is the determined cDNA sequence for LST-S1-A-12D  
SEQ ID NO: 85 is the determined cDNA sequence for LST-S1-A-11E  
SEQ ID NO: 86 is the determined cDNA sequence for LST-S1-A-12E  
SEQ ID NO: 87 is the determined cDNA sequence for L513S (T3).  
10 SEQ ID NO: 88 is the determined cDNA sequence for L513S contig 1.  
SEQ ID NO: 89 is a first determined cDNA sequence for L514S.  
SEQ ID NO: 90 is a second determined cDNA sequence for L514S.  
SEQ ID NO: 91 is a first determined cDNA sequence for L516S.  
SEQ ID NO: 92 is a second determined cDNA sequence for L516S.  
15 SEQ ID NO: 93 is the determined cDNA sequence for L517S.  
SEQ ID NO: 94 is the extended cDNA sequence for LST-S1-169 (also known as L519S).  
SEQ ID NO: 95 is a first determined cDNA sequence for L520S.  
SEQ ID NO: 96 is a second determined cDNA sequence for L520S.  
20 SEQ ID NO: 97 is a first determined cDNA sequence for L521S.  
SEQ ID NO: 98 is a second determined cDNA sequence for L521S.  
SEQ ID NO: 99 is the determined cDNA sequence for L522S.  
SEQ ID NO: 100 is the determined cDNA sequence for L523S.  
SEQ ID NO: 101 is the determined cDNA sequence for L524S.  
25 SEQ ID NO: 102 is the determined cDNA sequence for L525S.  
SEQ ID NO: 103 is the determined cDNA sequence for L526S.  
SEQ ID NO: 104 is the determined cDNA sequence for L527S.  
SEQ ID NO: 105 is the determined cDNA sequence for L528S.  
SEQ ID NO: 106 is the determined cDNA sequence for L529S.  
30 SEQ ID NO: 107 is a first determined cDNA sequence for L530S.



- SEQ ID NO: 108 is a second determined cDNA sequence for L530S.
- SEQ ID NO: 109 is the determined full-length cDNA sequence for L531S short form
- SEQ ID NO: 110 is the amino acid sequence encoded by SEQ ID NO: 109.
- SEQ ID NO: 111 is the determined full-length cDNA sequence for L531S long form
- 5 SEQ ID NO: 112 is the amino acid sequence encoded by SEQ ID NO: 111.
- SEQ ID NO: 113 is the determined full-length cDNA sequence for L520S.
- SEQ ID NO: 114 is the amino acid sequence encoded by SEQ ID NO: 113.
- SEQ ID NO: 115 is the determined cDNA sequence for contig 1.
- SEQ ID NO: 116 is the determined cDNA sequence for contig 3.
- 10 SEQ ID NO: 117 is the determined cDNA sequence for contig 4.
- SEQ ID NO: 118 is the determined cDNA sequence for contig 5.
- SEQ ID NO: 119 is the determined cDNA sequence for contig 7.
- SEQ ID NO: 120 is the determined cDNA sequence for contig 8.
- SEQ ID NO: 121 is the determined cDNA sequence for contig 9.
- 15 SEQ ID NO: 122 is the determined cDNA sequence for contig 10.
- SEQ ID NO: 123 is the determined cDNA sequence for contig 12.
- SEQ ID NO: 124 is the determined cDNA sequence for contig 11.
- SEQ ID NO: 125 is the determined cDNA sequence for contig 13 (also known as L761P).
- 20 SEQ ID NO: 126 is the determined cDNA sequence for contig 15.
- SEQ ID NO: 127 is the determined cDNA sequence for contig 16.
- SEQ ID NO: 128 is the determined cDNA sequence for contig 17.
- SEQ ID NO: 129 is the determined cDNA sequence for contig 19.
- SEQ ID NO: 130 is the determined cDNA sequence for contig 20.
- 25 SEQ ID NO: 131 is the determined cDNA sequence for contig 22.
- SEQ ID NO: 132 is the determined cDNA sequence for contig 24.
- SEQ ID NO: 133 is the determined cDNA sequence for contig 29.
- SEQ ID NO: 134 is the determined cDNA sequence for contig 31.
- SEQ ID NO: 135 is the determined cDNA sequence for contig 33.
- 30 SEQ ID NO: 136 is the determined cDNA sequence for contig 38.

- SEQ ID NO: 137 is the determined cDNA sequence for contig 39.  
SEQ ID NO: 138 is the determined cDNA sequence for contig 41.  
SEQ ID NO: 139 is the determined cDNA sequence for contig 43.  
SEQ ID NO: 140 is the determined cDNA sequence for contig 44.  
5 SEQ ID NO: 141 is the determined cDNA sequence for contig 45.  
SEQ ID NO: 142 is the determined cDNA sequence for contig 47.  
SEQ ID NO: 143 is the determined cDNA sequence for contig 48.  
SEQ ID NO: 144 is the determined cDNA sequence for contig 49.  
SEQ ID NO: 145 is the determined cDNA sequence for contig 50.  
10 SEQ ID NO: 146 is the determined cDNA sequence for contig 53.  
SEQ ID NO: 147 is the determined cDNA sequence for contig 54.  
SEQ ID NO: 148 is the determined cDNA sequence for contig 56.  
SEQ ID NO: 149 is the determined cDNA sequence for contig 57.  
SEQ ID NO: 150 is the determined cDNA sequence for contig 58.  
15 SEQ ID NO: 151 is the full-length cDNA sequence for L530S.  
SEQ ID NO: 152 is the amino acid sequence encoded by SEQ ID NO: 151  
SEQ ID NO: 153 is the full-length cDNA sequence of a first variant of L514S  
SEQ ID NO: 154 is the full-length cDNA sequence of a second variant of L514S  
SEQ ID NO: 155 is the amino acid sequence encoded by SEQ ID NO: 153.  
20 SEQ ID NO: 156 is the amino acid sequence encoded by SEQ ID NO: 154.  
SEQ ID NO: 157 is the determined cDNA sequence for contig 59.  
SEQ ID NO: 158 is the full-length cDNA sequence for L763P (also referred to as contig  
22).  
SEQ ID NO: 159 is the amino acid sequence encoded by SEQ ID NO: 158.  
25 SEQ ID NO: 160 is the full-length cDNA sequence for L762P (also referred to as contig  
17).  
SEQ ID NO: 161 is the amino acid sequence encoded by SEQ ID NO: 160.  
SEQ ID NO: 162 is the determined cDNA sequence for L515S.  
SEQ ID NO: 163 is the full-length cDNA sequence of a first variant of L524S.  
30 SEQ ID NO: 164 is the full-length cDNA sequence of a second variant of L524S.

- SEQ ID NO: 165 is the amino acid sequence encoded by SEQ ID NO: 163.  
SEQ ID NO: 166 is the amino acid sequence encoded by SEQ ID NO: 164.  
SEQ ID NO: 167 is the full-length cDNA sequence of a first variant of L762P.  
SEQ ID NO: 168 is the full-length cDNA sequence of a second variant of L762P.
- 5 SEQ ID NO: 169 is the amino acid sequence encoded by SEQ ID NO: 167.  
SEQ ID NO: 170 is the amino acid sequence encoded by SEQ ID NO: 168.  
SEQ ID NO: 171 is the full-length cDNA sequence for L773P (also referred to as contig 56).
- SEQ ID NO: 172 is the amino acid sequence encoded by SEQ ID NO: 171.
- 10 SEQ ID NO: 173 is an extended cDNA sequence for L519S.  
SEQ ID NO: 174 is the amino acid sequence encoded by SEQ ID NO: 174.  
SEQ ID NO: 175 is the full-length cDNA sequence for L523S.  
SEQ ID NO: 176 is the amino acid sequence encoded by SEQ ID NO: 175.  
SEQ ID NO: 177 is the determined cDNA sequence for LST-sub5-7A.
- 15 SEQ ID NO: 178 is the determined cDNA sequence for LST-sub5-8G.  
SEQ ID NO: 179 is the determined cDNA sequence for LST-sub5-8H.  
SEQ ID NO: 180 is the determined cDNA sequence for LST-sub5-10B.  
SEQ ID NO: 181 is the determined cDNA sequence for LST-sub5-10H.  
SEQ ID NO: 182 is the determined cDNA sequence for LST-sub5-12B.
- 20 SEQ ID NO: 183 is the determined cDNA sequence for LST-sub5-11C.  
SEQ ID NO: 184 is the determined cDNA sequence for LST-sub6-1c.  
SEQ ID NO: 185 is the determined cDNA sequence for LST-sub6-2f.  
SEQ ID NO: 186 is the determined cDNA sequence for LST-sub6-2G.  
SEQ ID NO: 187 is the determined cDNA sequence for LST-sub6-4d.
- 25 SEQ ID NO: 188 is the determined cDNA sequence for LST-sub6-4e.  
SEQ ID NO: 189 is the determined cDNA sequence for LST-sub6-4f.  
SEQ ID NO: 190 is the determined cDNA sequence for LST-sub6-3h.  
SEQ ID NO: 191 is the determined cDNA sequence for LST-sub6-5d.  
SEQ ID NO: 192 is the determined cDNA sequence for LST-sub6-5h.
- 30 SEQ ID NO: 193 is the determined cDNA sequence for LST-sub6-6h.

- SEQ ID NO: 194 is the determined cDNA sequence for LST-sub6-7a.  
SEQ ID NO: 195 is the determined cDNA sequence for LST-sub6-8a.  
SEQ ID NO: 196 is the determined cDNA sequence for LST-sub6-7d.  
SEQ ID NO: 197 is the determined cDNA sequence for LST-sub6-7e.  
5 SEQ ID NO: 198 is the determined cDNA sequence for LST-sub6-8e.  
SEQ ID NO: 199 is the determined cDNA sequence for LST-sub6-7g.  
SEQ ID NO: 200 is the determined cDNA sequence for LST-sub6-9f.  
SEQ ID NO: 201 is the determined cDNA sequence for LST-sub6-9h.  
SEQ ID NO: 202 is the determined cDNA sequence for LST-sub6-11b.  
10 SEQ ID NO: 203 is the determined cDNA sequence for LST-sub6-11c.  
SEQ ID NO: 204 is the determined cDNA sequence for LST-sub6-12c.  
SEQ ID NO: 205 is the determined cDNA sequence for LST-sub6-12e.  
SEQ ID NO: 206 is the determined cDNA sequence for LST-sub6-12f.  
SEQ ID NO: 207 is the determined cDNA sequence for LST-sub6-11g.  
15 SEQ ID NO: 208 is the determined cDNA sequence for LST-sub6-12g.  
SEQ ID NO: 209 is the determined cDNA sequence for LST-sub6-12h.  
SEQ ID NO: 210 is the determined cDNA sequence for LST-sub6-II-1a.  
SEQ ID NO: 211 is the determined cDNA sequence for LST-sub6-II-2b.  
SEQ ID NO: 212 is the determined cDNA sequence for LST-sub6-II-2g.  
20 SEQ ID NO: 213 is the determined cDNA sequence for LST-sub6-II-1h.  
SEQ ID NO: 214 is the determined cDNA sequence for LST-sub6-II-4a.  
SEQ ID NO: 215 is the determined cDNA sequence for LST-sub6-II-4b.  
SEQ ID NO: 216 is the determined cDNA sequence for LST-sub6-II-3e.  
SEQ ID NO: 217 is the determined cDNA sequence for LST-sub6-II-4f.  
25 SEQ ID NO: 218 is the determined cDNA sequence for LST-sub6-II-4g.  
SEQ ID NO: 219 is the determined cDNA sequence for LST-sub6-II-4h.  
SEQ ID NO: 220 is the determined cDNA sequence for LST-sub6-II-5c.  
SEQ ID NO: 221 is the determined cDNA sequence for LST-sub6-II-5e.  
SEQ ID NO: 222 is the determined cDNA sequence for LST-sub6-II-6f.  
30 SEQ ID NO: 223 is the determined cDNA sequence for LST-sub6-II-5g.

- SEQ ID NO: 224 is the determined cDNA sequence for LST-sub6-II-6g.
- SEQ ID NO: 225 is the amino acid sequence for L528S.
- SEQ ID NO: 226-251 are synthetic peptides derived from L762P.
- SEQ ID NO: 252 is the expressed amino acid sequence of L514S.
- 5 SEQ ID NO: 253 is the DNA sequence corresponding to SEQ ID NO: 252.
- SEQ ID NO: 254 is the DNA sequence of a L762P expression construct.
- SEQ ID NO: 255 is the determined cDNA sequence for clone 23785.
- SEQ ID NO: 256 is the determined cDNA sequence for clone 23786.
- SEQ ID NO: 257 is the determined cDNA sequence for clone 23788.
- 10 SEQ ID NO: 258 is the determined cDNA sequence for clone 23790.
- SEQ ID NO: 259 is the determined cDNA sequence for clone 23793.
- SEQ ID NO: 260 is the determined cDNA sequence for clone 23794.
- SEQ ID NO: 261 is the determined cDNA sequence for clone 23795.
- SEQ ID NO: 262 is the determined cDNA sequence for clone 23796.
- 15 SEQ ID NO: 263 is the determined cDNA sequence for clone 23797.
- SEQ ID NO: 264 is the determined cDNA sequence for clone 23798.
- SEQ ID NO: 265 is the determined cDNA sequence for clone 23799.
- SEQ ID NO: 266 is the determined cDNA sequence for clone 23800.
- SEQ ID NO: 267 is the determined cDNA sequence for clone 23802.
- 20 SEQ ID NO: 268 is the determined cDNA sequence for clone 23803.
- SEQ ID NO: 269 is the determined cDNA sequence for clone 23804.
- SEQ ID NO: 270 is the determined cDNA sequence for clone 23805.
- SEQ ID NO: 271 is the determined cDNA sequence for clone 23806.
- SEQ ID NO: 272 is the determined cDNA sequence for clone 23807.
- 25 SEQ ID NO: 273 is the determined cDNA sequence for clone 23808.
- SEQ ID NO: 274 is the determined cDNA sequence for clone 23809.
- SEQ ID NO: 275 is the determined cDNA sequence for clone 23810.
- SEQ ID NO: 276 is the determined cDNA sequence for clone 23811.
- SEQ ID NO: 277 is the determined cDNA sequence for clone 23812.
- 30 SEQ ID NO: 278 is the determined cDNA sequence for clone 23813.

- SEQ ID NO: 279 is the determined cDNA sequence for clone 23815.  
SEQ ID NO: 280 is the determined cDNA sequence for clone 25298.  
SEQ ID NO: 281 is the determined cDNA sequence for clone 25299.  
SEQ ID NO: 282 is the determined cDNA sequence for clone 25300.  
5 SEQ ID NO: 283 is the determined cDNA sequence for clone 25301  
SEQ ID NO: 284 is the determined cDNA sequence for clone 25304  
SEQ ID NO: 285 is the determined cDNA sequence for clone 25309.  
SEQ ID NO: 286 is the determined cDNA sequence for clone 25312.  
SEQ ID NO: 287 is the determined cDNA sequence for clone 25317.  
10 SEQ ID NO:288 is the determined cDNA sequence for clone 25321.  
SEQ ID NO:289 is the determined cDNA sequence for clone 25323.  
SEQ ID NO:290 is the determined cDNA sequence for clone 25327.  
SEQ ID NO:291 is the determined cDNA sequence for clone 25328.  
SEQ ID NO:292 is the determined cDNA sequence for clone 25332.  
15 SEQ ID NO:293 is the determined cDNA sequence for clone 25333.  
SEQ ID NO:294 is the determined cDNA sequence for clone 25336.  
SEQ ID NO:295 is the determined cDNA sequence for clone 25340.  
SEQ ID NO:296 is the determined cDNA sequence for clone 25342.  
SEQ ID NO:297 is the determined cDNA sequence for clone 25356.  
20 SEQ ID NO:298 is the determined cDNA sequence for clone 25357.  
SEQ ID NO:299 is the determined cDNA sequence for clone 25361.  
SEQ ID NO:300 is the determined cDNA sequence for clone 25363.  
SEQ ID NO:301 is the determined cDNA sequence for clone 25397.  
SEQ ID NO:302 is the determined cDNA sequence for clone 25402.  
25 SEQ ID NO:303 is the determined cDNA sequence for clone 25403.  
SEQ ID NO:304 is the determined cDNA sequence for clone 25405.  
SEQ ID NO:305 is the determined cDNA sequence for clone 25407.  
SEQ ID NO:306 is the determined cDNA sequence for clone 25409.  
SEQ ID NO:307 is the determined cDNA sequence for clone 25396.  
30 SEQ ID NO:308 is the determined cDNA sequence for clone 25414.

- SEQ ID NO:309 is the determined cDNA sequence for clone 25410.  
SEQ ID NO:310 is the determined cDNA sequence for clone 25406.  
SEQ ID NO:311 is the determined cDNA sequence for clone 25306.  
SEQ ID NO:312 is the determined cDNA sequence for clone 25362.  
5 SEQ ID NO:313 is the determined cDNA sequence for clone 25360.  
SEQ ID NO:314 is the determined cDNA sequence for clone 25398.  
SEQ ID NO:315 is the determined cDNA sequence for clone 25355.  
SEQ ID NO:316 is the determined cDNA sequence for clone 25351.  
SEQ ID NO:317 is the determined cDNA sequence for clone 25331.  
10 SEQ ID NO:318 is the determined cDNA sequence for clone 25338.  
SEQ ID NO:319 is the determined cDNA sequence for clone 25335.  
SEQ ID NO:320 is the determined cDNA sequence for clone 25329.  
SEQ ID NO:321 is the determined cDNA sequence for clone 25324.  
SEQ ID NO:322 is the determined cDNA sequence for clone 25322.  
15 SEQ ID NO:323 is the determined cDNA sequence for clone 25319.  
SEQ ID NO:324 is the determined cDNA sequence for clone 25316.  
SEQ ID NO:325 is the determined cDNA sequence for clone 25311.  
SEQ ID NO:326 is the determined cDNA sequence for clone 25310.  
SEQ ID NO:327 is the determined cDNA sequence for clone 25302.  
20 SEQ ID NO:328 is the determined cDNA sequence for clone 25315.  
SEQ ID NO:329 is the determined cDNA sequence for clone 25308.  
SEQ ID NO:330 is the determined cDNA sequence for clone 25303.  
SEQ ID NOs:331-337 are the cDNA sequences of isoforms of the p53 tumor suppressor homologue, p63 (also referred to as L530S).  
25 SEQ ID NOs:338-344 are the amino acid sequences encoded by SEQ ID NOs:331-337, respectively  
SEQ ID NO:345 is a second cDNA sequence for the antigen L763P.  
SEQ ID NO:346 is the amino acid sequence encoded by the sequence of SEQ ID NO: 345.  
30 SEQ ID NO:347 is a determined full-length cDNA sequence for L523S.

- SEQ ID NO:348 is the amino acid sequence encoded by SEQ ID NO: 347.
- SEQ ID NO:349 is the cDNA sequence encoding the N-terminal portion of L773P.
- SEQ ID NO:350 is the amino acid sequence of the N-terminal portion of L773P.
- SEQ ID NO:351 is the DNA sequence for a fusion of Ra12 and the N-terminal portion
- 5 of L763P
- SEQ ID NO:352 is the amino acid sequence of the fusion of Ra12 and the N-terminal portion of L763P
- SEQ ID NO:353 is the DNA sequence for a fusion of Ra12 and the C-terminal portion of L763P
- 10 SEQ ID NO:354 is the amino acid sequence of the fusion of Ra12 and the C-terminal portion of L763P
- SEQ ID NO:355 is a primer.
- SEQ ID NO:356 is a primer.
- SEQ ID NO:357 is the protein sequence of expressed recombinant L762P.
- 15 SEQ ID NO:358 is the DNA sequence of expressed recombinant L762P.
- SEQ ID NO:359 is a primer.
- SEQ ID NO:360 is a primer.
- SEQ ID NO:361 is the protein sequence of expressed recombinant L773P A.
- SEQ ID NO:362 is the DNA sequence of expressed recombinant L773P A.
- 20 SEQ ID NO:363 is an epitope derived from clone L773P polypeptide.
- SEQ ID NO:364 is a polynucleotide encoding the polypeptide of SEQ ID NO:363.
- SEQ ID NO:365 is an epitope derived from clone L773P polypeptide.
- SEQ ID NO:366 is a polynucleotide encoding the polypeptide of SEQ ID NO:365.
- SEQ ID NO:367 is an epitope consisting of amino acids 571-590 of SEQ ID NO:161,
- 25 clone L762P.
- SEQ ID NO:368 is the full-length DNA sequence for contig 13 (SEQ ID NO:125), also referred to as L761P.
- SEQ ID NO:369 is the protein sequence encoded by the DNA sequence of SEQ ID NO:368.
- 30 SEQ ID NO:370 is an L762P DNA sequence from nucleotides 2071-2130.



- SEQ ID NO:371 is an L762P DNA sequence from nucleotides 1441-1500.  
SEQ ID NO:372 is an L762P DNA sequence from nucleotides 1936-1955.  
SEQ ID NO:373 is an L762P DNA sequence from nucleotides 2620-2679.  
SEQ ID NO:374 is an L762P DNA sequence from nucleotides 1801-1860.
- 5 SEQ ID NO:375 is an L762P DNA sequence from nucleotides 1531-1591.  
SEQ ID NO:376 is the amino acid sequence of the L762P peptide encoded by SEQ ID NO:373.  
SEQ ID NO:377 is the amino acid sequence of the L762P peptide encoded by SEQ ID NO:370.
- 10 SEQ ID NO:378 is the amino acid sequence of the L762P peptide encoded by SEQ ID NO:372.  
SEQ ID NO:379 is the amino acid sequence of the L762P peptide encoded by SEQ ID NO:374.  
SEQ ID NO:380 is the amino acid sequence of the L762P peptide encoded by SEQ ID
- 15 NO:371.  
SEQ ID NO:381 is the amino acid sequence of the L762P peptide encoded by SEQ ID NO:375.  
SEQ ID NO:382 is the amino acid sequence of an epitope of L762P.  
SEQ ID NOs:383-386 are PCR primers.
- 20 SEQ ID NOs:387-395 are the amino acid sequences of L773P peptides.  
SEQ ID NOs:396-419 are the amino acid sequences of L523S peptides.  
SEQ ID NO:420 is the determined cDNA sequence for clone #19014.  
SEQ ID NO:421 is the forward primer PDM-278 for the L514S-13160 coding region.  
SEQ ID NO:422 is the reverse primer PDM-278 for the L514S-13160 coding region.
- 25 SEQ ID NO:423 is the amino acid sequence for the expressed recombinant L514S.  
SEQ ID NO:424 is the DNA coding sequence for the recombinant L514S.  
SEQ ID NO:425 is the forward primer PDM-414 for the L523S coding region.  
SEQ ID NO:426 is the reverse primer PDM-414 for the L523S coding region.  
SEQ ID NO:427 is the amino acid sequence for the expressed recombinant L523S.
- 30 SEQ ID NO:428 is the DNA coding sequence for the recombinant L523S.

- SEQ ID NO:429 is the reverse primer PDM-279 for the L762PA coding region.
- SEQ ID NO:430 is the amino acid sequence for the expressed recombinant L762PA.
- SEQ ID NO:431 is the DNA coding sequence for the recombinant L762PA.
- SEQ ID NO:432 is the reverse primer PDM-300 for the L773P coding region.
- 5 SEQ ID NO:433 is the amino acid sequence of the expressed recombinant L773P.
- SEQ ID NO:434 is the DNA coding sequence for the recombinant L773P.
- SEQ ID NO:435 is the forward primer for TCR Valpha8.
- SEQ ID NO:436 is the reverse primer for TCR Valpha8.
- SEQ ID NO:437 is the forward primer for TCR Vbeta8.
- 10 SEQ ID NO:438 is the reverse primer for TCR Vbeta8.
- SEQ ID NO:439 is the TCR Valpha DNA sequence of the TCR clone specific for the lung antigen L762P.
- SEQ ID NO:440 is the TCR Vbeta DNA sequence of the TCR clone specific for the lung antigen L762P.
- 15 SEQ ID NO:441 is the amino acid sequence of L763 peptide #2684.
- SEQ ID NO:442 is the predicted full-length cDNA for the cloned partial sequence of clone L529S (SEQ ID NO:106).
- SEQ ID NO:443 is the deduced amino acid sequence encoded by SEQ ID NO:442
- SEQ ID NO:444 is the forward primer PDM-734 for the coding region of clone L523S.
- 20 SEQ ID NO:445 is the reverse primer PDM-735 for the coding region of clone L523S.
- SEQ ID NO:446 is the amino acid sequence for the expressed recombinant L523S.
- SEQ ID NO:447 is the DNA coding sequence for the recombinant L523S.
- SEQ ID NO:448 is another forward primer PDM-733 for the coding region of clone L523S.
- 25 SEQ ID NO:449 is the amino acid sequence for a second expressed recombinant L523S.
- SEQ ID NO:450 is the DNA coding sequence for a second recombinant L523S.
- SEQ ID NO:451 corresponds to amino acids 86-110, an epitope of L514S-specific in the generation of antibodies.
- SEQ ID NO:452 corresponds to amino acids 21-45, an epitope of L514S-specific in the
- 30 generation of antibodies.

- SEQ ID NO:453 corresponds to amino acids 121-135, an epitope of L514S-specific in the generation of antibodies.
- SEQ ID NO:454 corresponds to amino acids 440-460, an epitope of L523S-specific in the generation of antibodies.
- 5 SEQ ID NO:455 corresponds to amino acids 156-175, an epitope of L523S-specific in the generation of antibodies.
- SEQ ID NO:456 corresponds to amino acids 326-345, an epitope of L523S-specific in the generation of antibodies.
- SEQ ID NO:457 corresponds to amino acids 40-59, an epitope of L523S-specific in the  
10 generation of antibodies.
- SEQ ID NO:458 corresponds to amino acids 80-99, an epitope of L523S-specific in the generation of antibodies.
- SEQ ID NO:459 corresponds to amino acids 160-179, an epitope of L523S-specific in the generation of antibodies.
- 15 SEQ ID NO:460 corresponds to amino acids 180-199, an epitope of L523S-specific in the generation of antibodies.
- SEQ ID NO:461 corresponds to amino acids 320-339, an epitope of L523S-specific in the generation of antibodies.
- SEQ ID NO:462 corresponds to amino acids 340-359, an epitope of L523S-specific in  
20 the generation of antibodies.
- SEQ ID NO:463 corresponds to amino acids 370-389, an epitope of L523S-specific in the generation of antibodies.
- SEQ ID NO:464 corresponds to amino acids 380-399, an epitope of L523S-specific in the generation of antibodies.
- 25 SEQ ID NO:465 corresponds to amino acids 37-55, an epitope of L523S-recognized by the L523S-specific CTL line 6B1.
- SEQ ID NO:466 corresponds to amino acids 41-51, the mapped antigenic epitope of L523S-recognized by the L523S-specific CTL line 6B1.
- SEQ ID NO:467 corresponds to the DNA sequence which encodes SEQ ID NO:466.

## DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed generally to compositions and their use in the therapy and diagnosis of cancer, particularly lung cancer. As described further below, illustrative compositions of the present invention include, but are not restricted to, polypeptides, particularly immunogenic polypeptides, polynucleotides encoding such polypeptides, antibodies and other binding agents, antigen presenting cells (APCs) and immune system cells (*e.g.*, T cells).

The practice of the present invention will employ, unless indicated specifically to the contrary, conventional methods of virology, immunology, microbiology, molecular biology and recombinant DNA techniques within the skill of the art, many of which are described below for the purpose of illustration. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, et al. *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); Maniatis et al. *Molecular Cloning: A Laboratory Manual* (1982); DNA Cloning: A Practical Approach, vol. I & II (D. Glover, ed.); *Oligonucleotide Synthesis* (N. Gait, ed., 1984); *Nucleic Acid Hybridization* (B. Hames & S. Higgins, eds., 1985); *Transcription and Translation* (B. Hames & S. Higgins, eds., 1984); *Animal Cell Culture* (R. Freshney, ed., 1986); Perbal, *A Practical Guide to Molecular Cloning* (1984).

All publications, patents and patent applications cited herein, whether *supra* or *infra*, are hereby incorporated by reference in their entirety.

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates otherwise.

## Polypeptide Compositions

As used herein, the term "polypeptide" is used in its conventional meaning, *i.e.*, as a sequence of amino acids. The polypeptides are not limited to a specific length of the product; thus, peptides, oligopeptides, and proteins are included

within the definition of polypeptide, and such terms may be used interchangeably herein unless specifically indicated otherwise. This term also does not refer to or exclude post-expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like, as well as other modifications known in the art, both naturally occurring and non-naturally occurring. A polypeptide may be an entire protein, or a subsequence thereof. Particular polypeptides of interest in the context of this invention are amino acid subsequences comprising epitopes, *i.e.*, antigenic determinants substantially responsible for the immunogenic properties of a polypeptide and being capable of evoking an immune response.

Particularly illustrative polypeptides of the present invention comprise those encoded by a polynucleotide sequence set forth in any one of SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467, or a sequence that hybridizes under moderately stringent conditions, or, alternatively, under highly stringent conditions, to a polynucleotide sequence set forth in any one of SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467. Certain illustrative polypeptides of the invention comprise amino acid sequences as set forth in any one of SEQ ID NOs:152, 155, 156, 165, 166, 169, 170, 172, 174, 176, 226-252, 338-344, 346, 350, 357, 361, 363, 365, 367, 369, 376-382, 387-419, 423, 427, 430, 433, 441, 443, 446, 449 and 451-466.

The polypeptides of the present invention are sometimes herein referred to as lung tumor proteins or lung tumor polypeptides, as an indication that their identification has been based at least in part upon their increased levels of expression in lung tumor samples. Thus, a "lung tumor polypeptide" or "lung tumor protein," refers

generally to a polypeptide sequence of the present invention, or a polynucleotide sequence encoding such a polypeptide, that is expressed in a substantial proportion of lung tumor samples, for example preferably greater than about 20%, more preferably greater than about 30%, and most preferably greater than about 50% or more of lung tumor samples tested, at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in normal tissues, as determined using a representative assay provided herein. A lung tumor polypeptide sequence of the invention, based upon its increased level of expression in tumor cells, has particular utility both as a diagnostic marker as well as a therapeutic target, as further described below.

In certain preferred embodiments, the polypeptides of the invention are immunogenic, *i.e.*, they react detectably within an immunoassay (such as an ELISA or T-cell stimulation assay) with antisera and/or T-cells from a patient with lung cancer. Screening for immunogenic activity can be performed using techniques well known to the skilled artisan. For example, such screens can be performed using methods such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In one illustrative example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, <sup>125</sup>I-labeled Protein A.

As would be recognized by the skilled artisan, immunogenic portions of the polypeptides disclosed herein are also encompassed by the present invention. An "immunogenic portion," as used herein, is a fragment of an immunogenic polypeptide of the invention that itself is immunologically reactive (*i.e.*, specifically binds) with B-cells and/or T-cell surface antigen receptors that recognize the polypeptide. Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they

specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well-known techniques.

In one preferred embodiment, an immunogenic portion of a polypeptide  
5 of the present invention is a portion that reacts with antisera and/or T-cells at a level that is not substantially less than the reactivity of the full-length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Preferably, the level of immunogenic activity of the immunogenic portion is at least about 50%, preferably at least about 70% and most preferably greater than about 90% of the immunogenicity for the full-length  
10 polypeptide. In some instances, preferred immunogenic portions will be identified that have a level of immunogenic activity greater than that of the corresponding full-length polypeptide, *e.g.*, having greater than about 100% or 150% or more immunogenic activity.

In certain other embodiments, illustrative immunogenic portions may  
15 include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other illustrative immunogenic portions will contain a small N- and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

In another embodiment, a polypeptide composition of the invention may  
20 also comprise one or more polypeptides that are immunologically reactive with T cells and/or antibodies generated against a polypeptide of the invention, particularly a polypeptide having an amino acid sequence disclosed herein, or to an immunogenic fragment or variant thereof.

In another embodiment of the invention, polypeptides are provided that  
25 comprise one or more polypeptides that are capable of eliciting T cells and/or antibodies that are immunologically reactive with one or more polypeptides described herein, or one or more polypeptides encoded by contiguous nucleic acid sequences contained in the polynucleotide sequences disclosed herein, or immunogenic fragments or variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of  
30 these sequences under conditions of moderate to high stringency.

The present invention, in another aspect, provides polypeptide fragments comprising at least about 5, 10, 15, 20, 25, 50, or 100 contiguous amino acids, or more, including all intermediate lengths, of a polypeptide compositions set forth herein, such as those set forth in SEQ ID NOs:152, 155, 156, 165, 166, 169, 170, 172, 174, 176, 5 226-252, 338-344, 346, 350, 357, 361, 363, 365, 367, 369, 376-382 and 387-419, 441, 443, 446, 449 and 451-466, or those encoded by a polynucleotide sequence set forth in a sequence of SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 10 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467.

In another aspect, the present invention provides variants of the polypeptide compositions described herein. Polypeptide variants generally encompassed by the present invention will typically exhibit at least about 70%, 75%, 15 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described below), along its length, to a polypeptide sequences set forth herein.

In one preferred embodiment, the polypeptide fragments and variants provide by the present invention are immunologically reactive with an antibody and/or 20 T-cell that reacts with a full-length polypeptide specifically set forth herein.

In another preferred embodiment, the polypeptide fragments and variants provided by the present invention exhibit a level of immunogenic activity of at least about 50%, preferably at least about 70%, and most preferably at least about 90% or more of that exhibited by a full-length polypeptide sequence specifically set forth 25 herein.

A polypeptide "variant," as the term is used herein, is a polypeptide that typically differs from a polypeptide specifically disclosed herein in one or more substitutions, deletions, additions and/or insertions. Such variants may be naturally occurring or may be synthetically generated, for example, by modifying one or more of 30 the above polypeptide sequences of the invention and evaluating their immunogenic



activity as described herein and/or using any of a number of techniques well known in the art.

For example, certain illustrative variants of the polypeptides of the invention include those in which one or more portions, such as an N-terminal leader  
5 sequence or transmembrane domain, have been removed. Other illustrative variants include variants in which a small portion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

In many instances, a variant will contain conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another  
10 amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. As described above, modifications may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a variant or derivative polypeptide  
15 with desirable characteristics, *e.g.*, with immunogenic characteristics. When it is desired to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, immunogenic variant or portion of a polypeptide of the invention, one skilled in the art will typically change one or more of the codons of the encoding DNA sequence according to Table 1.

For example, certain amino acids may be substituted for other amino  
20 acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence  
25 substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

30

TABLE 1

Amino Acids			Codons						
Alanine	Ala	A	GCA	GCC	GCG	GCU			
Cysteine	Cys	C	UGC	UGU					
Aspartic acid	Asp	D	GAC	GAU					
Glutamic acid	Glu	E	GAA	GAG					
Phenylalanine	Phe	F	UUC	UUU					
Glycine	Gly	G	GGA	GGC	GGG	GGU			
Histidine	His	H	CAC	CAU					
Isoleucine	Ile	I	AUA	AUC	AUU				
Lysine	Lys	K	AAA	AAG					
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU	
Methionine	Met	M	AUG						
Asparagine	Asn	N	AAC	AAU					
Proline	Pro	P	CCA	CCC	CCG	CCU			
Glutamine	Gln	Q	CAA	CAG					
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU	
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU	
Threonine	Thr	T	ACA	ACC	ACG	ACU			
Valine	Val	V	GUA	GUC	GUG	GUU			
Tryptophan	Trp	W	UGG						
Tyrosine	Tyr	Y	UAC	UAU					

In making such changes, the hydropathic index of amino acids may be  
 5 considered. The importance of the hydropathic amino acid index in conferring  
 interactive biologic function on a protein is generally understood in the art (Kyte and  
 Doolittle, 1982, incorporated herein by reference). It is accepted that the relative  
 hydropathic character of the amino acid contributes to the secondary structure of the  
 resultant protein, which in turn defines the interaction of the protein with other  
 10 molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and  
 the like. Each amino acid has been assigned a hydropathic index on the basis of its

hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5);  
5 glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within  $\pm 2$  is preferred, those within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of  
15 its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0  $\pm$  1); glutamate (+3.0  $\pm$  1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5  $\pm$  1); alanine (-0.5); histidine (-0.5); cysteine  
20 (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within  $\pm 2$   
25 is preferred, those within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that  
30 take various of the foregoing characteristics into consideration are well known to those

of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of  
5 flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl-methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

10 Amino acid substitutions may further be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values  
15 include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain non-conservative changes. In a  
20 preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example; the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

25 As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For  
30 example, a polypeptide may be conjugated to an immunoglobulin Fc region.

When comparing polypeptide sequences, two sequences are said to be "identical" if the sequence of amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

- Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA* 80:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics

Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment.

In one preferred approach, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Within other illustrative embodiments, a polypeptide may be a fusion polypeptide that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known

tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the polypeptide or to enable the polypeptide to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the polypeptide.

Fusion polypeptides may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion polypeptide is expressed as a recombinant polypeptide, allowing the production of increased levels, relative to a non-fused polypeptide, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion polypeptide that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion polypeptide using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene* 40:39-46, 1985; Murphy et al.,

*Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to  
5 separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and  
10 transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

The fusion polypeptide can comprise a polypeptide as described herein together with an unrelated immunogenic protein, such as an immunogenic protein capable of eliciting a recall response. Examples of such proteins include tetanus,  
15 tuberculosis and hepatitis proteins (*see*, for example, Stoute et al. *New Engl. J. Med.*, 336:86-91, 1997).

In one preferred embodiment, the immunological fusion partner is derived from a *Mycobacterium* sp., such as a *Mycobacterium tuberculosis*-derived Ra12 fragment. Ra12 compositions and methods for their use in enhancing the expression  
20 and/or immunogenicity of heterologous polynucleotide/polypeptide sequences is described in U.S. Patent Application 60/158,585, the disclosure of which is incorporated herein by reference in its entirety. Briefly, Ra12 refers to a polynucleotide region that is a subsequence of a *Mycobacterium tuberculosis* MTB32A nucleic acid. MTB32A is a serine protease of 32 KD molecular weight encoded by a gene in virulent  
25 and avirulent strains of *M. tuberculosis*. The nucleotide sequence and amino acid sequence of MTB32A have been described (for example, U.S. Patent Application 60/158,585; *see also*, Skeiky et al., *Infection and Immunity*. (1999) 67:3998-4007, incorporated herein by reference). C-terminal fragments of the MTB32A coding sequence express at high levels and remain as a soluble polypeptides throughout the  
30 purification process. Moreover, Ra12 may enhance the immunogenicity of heterologous



immunogenic polypeptides with which it is fused. One preferred Ra12 fusion polypeptide comprises a 14 KD C-terminal fragment corresponding to amino acid residues 192 to 323 of MTB32A.

Other preferred Ra12 polynucleotides generally comprise at least about  
5 15 consecutive nucleotides, at least about 30 nucleotides, at least about 60 nucleotides, at least about 100 nucleotides, at least about 200 nucleotides, or at least about 300 nucleotides that encode a portion of a Ra12 polypeptide.

Ra12 polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a Ra12 polypeptide or a portion thereof) or may  
10 comprise a variant of such a sequence. Ra12 polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the biological activity of the encoded fusion polypeptide is not substantially diminished, relative to a fusion polypeptide comprising a native Ra12 polypeptide. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most  
15 preferably at least about 90% identity to a polynucleotide sequence that encodes a native Ra12 polypeptide or a portion thereof.

Within other preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenza B* (WO 91/18926). Preferably, a protein D derivative comprises  
20 approximately the first third of the protein (*e.g.*, the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer).  
25 The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein  
30 known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is

derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the *LytA* gene; *Gene* 43:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (see *Biotechnology* 10:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion polypeptide. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

Yet another illustrative embodiment involves fusion polypeptides, and the polynucleotides encoding them, wherein the fusion partner comprises a targeting signal capable of directing a polypeptide to the endosomal/lysosomal compartment, as described in U.S. Patent No. 5,633,234. An immunogenic polypeptide of the invention, when fused with this targeting signal, will associate more efficiently with MHC class II molecules and thereby provide enhanced in vivo stimulation of CD4<sup>+</sup> T-cells specific for the polypeptide.

Polypeptides of the invention are prepared using any of a variety of well known synthetic and/or recombinant techniques, the latter of which are further described below. Polypeptides, portions and other variants generally less than about 150 amino acids can be generated by synthetic means, using techniques well known to those of ordinary skill in the art. In one illustrative example, such polypeptides are synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. See Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

In general, polypeptide compositions (including fusion polypeptides) of the invention are isolated. An "isolated" polypeptide is one that is removed from its original environment. For example, a naturally-occurring protein or polypeptide is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are also purified, *e.g.*, are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure.

### Polynucleotide Compositions

The present invention, in other aspects, provides polynucleotide compositions. The terms "DNA" and "polynucleotide" are used essentially interchangeably herein to refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. "Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA molecule does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA molecule as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be understood by those skilled in the art, the polynucleotide compositions of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

As will be also recognized by the skilled artisan, polynucleotides of the invention may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules may include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the

present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a polypeptide/protein of the invention or a portion thereof) or  
5 may comprise a sequence that encodes a variant or derivative, preferably and immunogenic variant or derivative, of such a sequence.

Therefore, according to another aspect of the present invention, polynucleotide compositions are provided that comprise some or all of a polynucleotide sequence set forth in any one of SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49,  
10 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467, complements of a polynucleotide sequence set forth in any one  
15 of SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467, and degenerate  
20 variants of a polynucleotide sequence set forth in any one of SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467. In certain preferred embodiments, the polynucleotide sequences set forth herein encode immunogenic polypeptides, as described above.

In other related embodiments, the present invention provides polynucleotide variants having substantial identity to the sequences disclosed herein in  
30 SEQ ID NOs:1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71,

73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347, 349, 358, 362, 364, 365, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467, for example those  
5 comprising at least 70% sequence identity, preferably at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide sequence of this invention using the methods described herein, (*e.g.*, BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine  
10 corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

Typically, polynucleotide variants will contain one or more substitutions, additions, deletions and/or insertions, preferably such that the immunogenicity of the  
15 polypeptide encoded by the variant polynucleotide is not substantially diminished relative to a polypeptide encoded by a polynucleotide sequence specifically set forth herein). The term "variants" should also be understood to encompass homologous genes of xenogenic origin.

In additional embodiments, the present invention provides  
20 polynucleotide fragments comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about 10, 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all  
25 intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

In another embodiment of the invention, polynucleotide compositions are provided that are capable of hybridizing under moderate to high stringency conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-60°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS. One skilled in the art will understand that the stringency of hybridization can be readily manipulated, such as by altering the salt content of the hybridization solution and/or the temperature at which the hybridization is performed. For example, in another embodiment, suitable highly stringent hybridization conditions include those described above, with the exception that the temperature of hybridization is increased, *e.g.*, to 60-65°C or 65-70°C.

In certain preferred embodiments, the polynucleotides described above, *e.g.*, polynucleotide variants, fragments and hybridizing sequences, encode polypeptides that are immunologically cross-reactive with a polypeptide sequence specifically set forth herein. In other preferred embodiments, such polynucleotides encode polypeptides that have a level of immunogenic activity of at least about 50%, preferably at least about 70%, and more preferably at least about 90% of that for a polypeptide sequence specifically set forth herein.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative polynucleotide segments with total lengths of about 10,000, about

5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

When comparing polynucleotide sequences, two sequences are said to be  
 5 "identical" if the sequence of nucleotides in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions,  
 10 usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR,  
 15 Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990)  
 20 Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and*  
 25 *Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad. Sci. USA* 80:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J.*  
 30 *Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988)

*Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

- 5                   One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent
- 10                   sequence identity for the polynucleotides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). Extension of
- 15                   the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for
- 20                   nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

- Preferably, the "percentage of sequence identity" is determined by
- 25                   comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The
- 30                   percentage is calculated by determining the number of positions at which the identical



nucleic acid bases occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

5           It will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present  
10 invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard  
15 techniques (such as hybridization, amplification and/or database sequence comparison).

          Therefore, in another embodiment of the invention, a mutagenesis approach, such as site-specific mutagenesis, is employed for the preparation of immunogenic variants and/or derivatives of the polypeptides described herein. By this approach, specific modifications in a polypeptide sequence can be made through  
20 mutagenesis of the underlying polynucleotides that encode them. These techniques provides a straightforward approach to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the polynucleotide.

          Site-specific mutagenesis allows the production of mutants through the  
25 use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise

change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or  
5 more properties of the encoded polypeptide, such as the immunogenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25  
10 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis  
15 include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is  
20 performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I  
25 Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term "oligonucleotide directed mutagenesis procedure" refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term "oligonucleotide directed mutagenesis procedure" is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

In another approach for the production of polypeptide variants of the present invention, recursive sequence recombination, as described in U.S. Patent No. 5,837,458, may be employed. In this approach, iterative cycles of recombination and screening or selection are performed to "evolve" individual polynucleotide variants of the invention having, for example, enhanced immunogenic activity.

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic

acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or  
5 complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of  
10 complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides  
15 or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the  
20 complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length  
25 complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the  
30 hybrid, and thereby improve the quality and degree of specific hybrid molecules

obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequences set forth herein, or to any continuous portion of the sequences, from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR<sup>TM</sup> technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

According to another embodiment of the present invention, polynucleotide compositions comprising antisense oligonucleotides are provided. Antisense oligonucleotides have been demonstrated to be effective and targeted inhibitors of protein synthesis, and, consequently, provide a therapeutic approach by which a disease can be treated by inhibiting the synthesis of proteins that contribute to the disease. The efficacy of antisense oligonucleotides for inhibiting protein synthesis is well established. For example, the synthesis of polygalacturonase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829). Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA<sub>A</sub> receptor and human EGF (Jaskulski *et al.*, Science. 1988 Jun 10;240(4858):1544-6; Vasanthakumar and Ahmed, Cancer Commun. 1989;1(4):225-32; Peris *et al.*, Brain Res Mol Brain Res. 1998 Jun 15;57(2):310-20; U. S. Patent 5,801,154; U.S. Patent 5,789,573; U. S. Patent 5,718,709 and U.S. Patent 5,610,288). Antisense constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683).

Therefore, in certain embodiments, the present invention provides oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein. Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence and determination of secondary structure,  $T_m$ , binding energy, and relative stability. Antisense compositions may be selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell. Highly preferred target regions of the mRNA, are those which are at or near the AUG translation initiation codon, and those sequences which are substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations can be performed, for example, using v.4 of the OLIGO primer analysis software and/or the BLASTN 2.0.5 algorithm software (Altschul *et al.*, Nucleic Acids Res. 1997, 25(17):3389-402).

The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, Nucleic Acids Res. 1997 Jul 15;25(14):2730-6). It has been demonstrated that several molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%).

Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane.

According to another embodiment of the invention, the polynucleotide compositions described herein are used in the design and preparation of ribozyme molecules for inhibiting expression of the tumor polypeptides and proteins of the present invention in tumor cells. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, *Proc Natl Acad Sci U S A.* 1987 Dec;84(24):8788-92; Forster and Symons, *Cell.* 1987 Apr 24;49(2):211-20). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et al.*, *Cell.* 1981 Dec;27(3 Pt 2):487-96; Michel and Westhof, *J Mol Biol.* 1990 Dec 5;216(3):585-610; Reinhold-Hurek and Shub, *Nature.* 1992 May 14;357(6374):173-6). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds



to a nucleic acid target to block its translation) since the concentration of ribozyme necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, Proc Natl Acad Sci U S A. 1992 Aug 15;89(16):7305-9). Thus, the specificity of action of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis  $\delta$  virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are described by Rossi *et al.* Nucleic Acids Res. 1992 Sep 11;20(17):4559-65. Examples of hairpin motifs are described by Hampel *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz, Biochemistry 1989 Jun 13;28(12):4929-33; Hampel *et al.*, Nucleic Acids Res. 1990 Jan 25;18(2):299-304 and U. S. Patent 5,631,359. An example of the hepatitis  $\delta$  virus motif is described by Perrotta and Been, Biochemistry. 1992 Dec 1;31(47):11843-52; an example of the RNaseP motif is described by Guerrier-Takada *et al.*, Cell. 1983 Dec;35(3 Pt 2):849-57; Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, Cell. 1990 May 18;61(4):685-96; Saville and Collins, Proc Natl Acad Sci U S A. 1991 Oct 1;88(19):8826-30; Collins and Olive, Biochemistry. 1993 Mar 23;32(11):2795-9); and an example of the Group I intron is described in (U. S. Patent 4,987,071). All that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart an

RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stint. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO

94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression  
5 vector. Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby.  
10 Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells. Ribozymes expressed from such promoters have been shown to function in mammalian cells. Such transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA  
15 vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

In another embodiment of the invention, peptide nucleic acids (PNAs) compositions are provided. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and Nielsen, *Antisense Nucleic Acid Drug*  
20 *Dev.* 1997 7(4) 431-37). PNA is able to be utilized in a number of methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (*Trends Biotechnol* 1997  
25 Jun;15(6):224-9). As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter, decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, *Science* 1991 Dec 6;254(5037):1497-500; Hanvey *et al.*, *Science*. 1992 Nov 27;258(5087):1481-5; Hyrup and Nielsen, *Bioorg Med Chem*. 1996 Jan;4(1):5-23). This chemistry has three important  
5 consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc or Fmoc protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used.

10 PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton *et al.*, *Bioorg Med Chem*. 1995 Apr;3(4):437-45). The manual protocol lends itself to the production of chemically modified PNAs or the simultaneous synthesis of families of  
15 closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this  
20 difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography, providing yields and purity of product similar to those observed during the synthesis of peptides.

25 Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or  
30 for specific functional requirements. Once synthesized, the identity of PNAs and their

- derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (for example, Norton *et al.*, Bioorg Med Chem. 1995 Apr;3(4):437-45; Petersen *et al.*, J Pept Sci. 1995 May-Jun;1(3):175-83; Orum *et al.*, Biotechniques. 1995 Sep;19(3):472-80; Footer *et al.*, Biochemistry. 1996 Aug 20;35(33):10673-9; Griffith *et al.*, Nucleic Acids Res. 1995 Aug 11;23(15):3003-8; Pardridge *et al.*, Proc Natl Acad Sci U S A. 1995 Jun 6;92(12):5592-6; Boffa *et al.*, Proc Natl Acad Sci U S A. 1995 Mar 14;92(6):1901-5; Gambacorti-Passerini *et al.*, Blood. 1996 Aug 15;88(4):1411-7; Armitage *et al.*, Proc Natl Acad Sci U S A. 1997 Nov 11;94(23):12320-5; Seeger *et al.*, Biotechniques. 1997 Sep;23(3):512-7). U.S. Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

- Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (Anal Chem. 1993 Dec 15;65(24):3545-9) and Jensen *et al.* (Biochemistry. 1997 Apr 22;36(16):5072-7). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen *et al.* using BIAcore™ technology.

- Other applications of PNAs that have been described and will be apparent to the skilled artisan include use in DNA strand invasion, antisense inhibition, mutational analysis, enhancers of transcription, nucleic acid purification, isolation of transcriptionally active genes, blocking of transcription factor binding, genome cleavage, biosensors, *in situ* hybridization, and the like.

#### Polynucleotide Identification, Characterization and Expression

- Polynucleotides compositions of the present invention may be identified, prepared and/or manipulated using any of a variety of well established techniques (see generally, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989, and other like references). For example, a polynucleotide may be identified, as described in more detail below, by

screening a microarray of cDNAs for tumor-associated expression (*i.e.*, expression that is at least two fold greater in a tumor than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using the microarray technology of Affymetrix, Inc. (Santa Clara, CA) according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as tumor cells.

Many template dependent processes are available to amplify a target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCR™) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCR™, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCR™ amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Any of a number of other template dependent processes, many of which are variations of the PCR™ amplification technique, are readily known and available in the art. Illustratively, some such methods include the ligase chain reaction (referred to as LCR), described, for example, in Eur. Pat. Appl. Publ. No. 320,308 and U.S. Patent No. 4,883,750; Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880; Strand Displacement Amplification (SDA) and Repair Chain

Reaction (RCR). Still other amplification methods are described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025. Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (PCT Intl. Pat. Appl. Publ. No. WO 88/10315), including nucleic acid sequence  
5 based amplification (NASBA) and 3SR. Eur. Pat. Appl. Publ. No. 329,822 describes a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA). PCT Intl. Pat. Appl. Publ. No. WO 89/06700 describes a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA  
10 ("ssDNA") followed by transcription of many RNA copies of the sequence. Other amplification methods such as "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) are also well-known to those of skill in the art.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (e.g., a tumor cDNA library)  
15 using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

20 For hybridization techniques, a partial sequence may be labeled (e.g., by nick-translation or end-labeling with <sup>32</sup>P) using well known techniques. A bacterial or bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (see Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor  
25 Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The  
30 complete sequence may then be determined using standard techniques, which may

involve generating a series of deletion clones. The resulting overlapping sequences can then assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

- Alternatively, amplification techniques, such as those described above,
- 5 can be useful for obtaining a full length coding sequence from a partial cDNA sequence. One such amplification technique is inverse PCR (*see* Triglia et al., *Nucl. Acids Res.* 16:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region.
- 10 Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate
- 15 extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et
- 20 al., *PCR Methods Applic. 1*:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids. Res.* 19:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

- In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as
- 25 that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (*e.g.*, NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.

- In other embodiments of the invention, polynucleotide sequences or
- 30 fragments thereof which encode polypeptides of the invention, or fusion proteins or



functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may  
5 be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression  
10 or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide  
15 encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction  
20 sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of  
25 polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical  
5 methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. et al. (1995) *Science* 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

10 A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) *Proteins, Structures and Molecular Principles*, WH Freeman and Co., New York, N.Y.) or other comparable techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation  
15 procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate  
20 expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA  
25 techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. et al. (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; 5 insect cell systems infected with virus expression vectors (*e.g.*, baculovirus); plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an 10 expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. 15 For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple 20 copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV may be advantageously used with an appropriate selectable marker.

In bacterial systems, any of a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors 25 which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of  $\beta$ -galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. 30

M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

An insect system may also be used to express a polypeptide of interest. For example, in one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence

will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. et al. (1994) *Proc. Natl. Acad. Sci.* 91 :3224-3227).

5           In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used  
10 to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

          Specific initiation signals may also be used to achieve more efficient  
15 translation of sequences encoding a polypeptide of interest. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion  
20 thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are  
25 appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162).

          In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to,  
30 acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation.

Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, COS, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1990) *Cell* 22:817-23) genes which can be employed in tk.sup.- or apt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol.* 150:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). The use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate

luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. et al. (1995) *Methods Mol. Biol.* 55:121-131).

- Although the presence/absence of marker gene expression suggests that
- 5 the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter.
- 10 Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

- Alternatively, host cells that contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-
- 15 RNA hybridizations and protein bioassay or immunoassay techniques which include, for example, membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

- A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies
- 20 specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed.
- 25 These and other assays are described, among other places, in Hampton, R. et al. (1990; *Serological Methods, a Laboratory Manual*, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; *J. Exp. Med.* 158:1211-1216).

- A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means
- 30 for producing labeled hybridization or PCR probes for detecting sequences related to

polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. et al. (1992, *Prot. Exp. Purif.* 3:263-281) while the enterokinase



cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; *DNA Cell Biol.* 12:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

#### Antibody Compositions, Fragments Thereof and Other Binding Agents

According to another aspect, the present invention further provides binding agents, such as antibodies and antigen-binding fragments thereof, that exhibit immunological binding to a tumor polypeptide disclosed herein, or to a portion, variant or derivative thereof. An antibody, or antigen-binding fragment thereof, is said to "specifically bind," "immunologically bind," and/or is "immunologically reactive" to a polypeptide of the invention if it reacts at a detectable level (within, for example, an ELISA assay) with the polypeptide, and does not react detectably with unrelated polypeptides under similar conditions.

Immunological binding, as used in this context, generally refers to the non-covalent interactions of the type which occur between an immunoglobulin molecule and an antigen for which the immunoglobulin is specific. The strength, or affinity of immunological binding interactions can be expressed in terms of the dissociation constant ( $K_d$ ) of the interaction, wherein a smaller  $K_d$  represents a greater affinity. Immunological binding properties of selected polypeptides can be quantified using methods well known in the art. One such method entails measuring the rates of antigen-binding site/antigen complex formation and dissociation, wherein those rates depend on the concentrations of the complex partners, the affinity of the interaction, and

on geometric parameters that equally influence the rate in both directions. Thus, both the "on rate constant" ( $K_{on}$ ) and the "off rate constant" ( $K_{off}$ ) can be determined by calculation of the concentrations and the actual rates of association and dissociation. The ratio of  $K_{off}/K_{on}$  enables cancellation of all parameters not related to affinity, and is thus equal to the dissociation constant  $K_d$ . See, generally, Davies et al. (1990) Annual Rev. Biochem. 59:439-473.

An "antigen-binding site," or "binding portion" of an antibody refers to the part of the immunoglobulin molecule that participates in antigen binding. The antigen binding site is formed by amino acid residues of the N-terminal variable ("V") regions of the heavy ("H") and light ("L") chains. Three highly divergent stretches within the V regions of the heavy and light chains are referred to as "hypervariable regions" which are interposed between more conserved flanking stretches known as "framework regions," or "FRs". Thus the term "FR" refers to amino acid sequences which are naturally found between and adjacent to hypervariable regions in immunoglobulins. In an antibody molecule, the three hypervariable regions of a light chain and the three hypervariable regions of a heavy chain are disposed relative to each other in three dimensional space to form an antigen-binding surface. The antigen-binding surface is complementary to the three-dimensional surface of a bound antigen, and the three hypervariable regions of each of the heavy and light chains are referred to as "complementarity-determining regions," or "CDRs."

Binding agents may be further capable of differentiating between patients with and without a cancer, such as lung cancer, using the representative assays provided herein. For example, antibodies or other binding agents that bind to a tumor protein will preferably generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, more preferably at least about 30% of patients. Alternatively, or in addition, the antibody will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein

for the presence of polypeptides that bind to the binding agent. Preferably, a statistically significant number of samples with and without the disease will be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve  
5 sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of  
10 a variety of techniques known to those of ordinary skill in the art. See, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into  
15 suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (e.g., mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier  
20 protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a  
25 suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the  
30 desired specificity (i.e., reactivity with the polypeptide of interest). Such cell lines may

be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells  
5 and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture  
10 supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable  
15 vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

20 A number of therapeutically useful molecules are known in the art which comprise antigen-binding sites that are capable of exhibiting immunological binding properties of an antibody molecule. The proteolytic enzyme papain preferentially cleaves IgG molecules to yield several fragments, two of which (the "F(ab)" fragments) each comprise a covalent heterodimer that includes an intact antigen-binding site. The  
25 enzyme pepsin is able to cleave IgG molecules to provide several fragments, including the "F(ab')<sub>2</sub>" fragment which comprises both antigen-binding sites. An "Fv" fragment can be produced by preferential proteolytic cleavage of an IgM, and on rare occasions IgG or IgA immunoglobulin molecule. Fv fragments are, however, more commonly derived using recombinant techniques known in the art. The Fv fragment includes a  
30 non-covalent V<sub>H</sub>::V<sub>L</sub> heterodimer including an antigen-binding site which retains much

of the antigen recognition and binding capabilities of the native antibody molecule. Inbar et al. (1972) Proc. Nat. Acad. Sci. USA 69:2659-2662; Hochman et al. (1976) Biochem 15:2706-2710; and Ehrlich et al. (1980) Biochem 19:4091-4096.

A single chain Fv ("sFv") polypeptide is a covalently linked  $V_H::V_L$  heterodimer which is expressed from a gene fusion including  $V_H$ - and  $V_L$ -encoding genes linked by a peptide-encoding linker. Huston et al. (1988) Proc. Nat. Acad. Sci. USA 85(16):5879-5883. A number of methods have been described to discern chemical structures for converting the naturally aggregated--but chemically separated--light and heavy polypeptide chains from an antibody V region into an sFv molecule which will fold into a three dimensional structure substantially similar to the structure of an antigen-binding site. See, e.g., U.S. Pat. Nos. 5,091,513 and 5,132,405, to Huston et al.; and U.S. Pat. No. 4,946,778, to Ladner et al.

Each of the above-described molecules includes a heavy chain and a light chain CDR set, respectively interposed between a heavy chain and a light chain FR set which provide support to the CDRs and define the spatial relationship of the CDRs relative to each other. As used herein, the term "CDR set" refers to the three hypervariable regions of a heavy or light chain V region. Proceeding from the N-terminus of a heavy or light chain, these regions are denoted as "CDR1," "CDR2," and "CDR3" respectively. An antigen-binding site, therefore, includes six CDRs, comprising the CDR set from each of a heavy and a light chain V region. A polypeptide comprising a single CDR, (e.g., a CDR1, CDR2 or CDR3) is referred to herein as a "molecular recognition unit." Crystallographic analysis of a number of antigen-antibody complexes has demonstrated that the amino acid residues of CDRs form extensive contact with bound antigen, wherein the most extensive antigen contact is with the heavy chain CDR3. Thus, the molecular recognition units are primarily responsible for the specificity of an antigen-binding site.

As used herein, the term "FR set" refers to the four flanking amino acid sequences which frame the CDRs of a CDR set of a heavy or light chain V region. Some FR residues may contact bound antigen; however, FRs are primarily responsible for folding the V region into the antigen-binding site, particularly the FR residues

directly adjacent to the CDRS. Within FRs, certain amino residues and certain structural features are very highly conserved. In this regard, all V region sequences contain an internal disulfide loop of around 90 amino acid residues. When the V regions fold into a binding-site, the CDRs are displayed as projecting loop motifs which form an antigen-binding surface. It is generally recognized that there are conserved structural regions of FRs which influence the folded shape of the CDR loops into certain "canonical" structures--regardless of the precise CDR amino acid sequence. Further, certain FR residues are known to participate in non-covalent interdomain contacts which stabilize the interaction of the antibody heavy and light chains.

A number of "humanized" antibody molecules comprising an antigen-binding site derived from a non-human immunoglobulin have been described, including chimeric antibodies having rodent V regions and their associated CDRs fused to human constant domains (Winter et al. (1991) *Nature* 349:293-299; Lobuglio et al. (1989) *Proc. Nat. Acad. Sci. USA* 86:4220-4224; Shaw et al. (1987) *J Immunol.* 138:4534-4538; and Brown et al. (1987) *Cancer Res.* 47:3577-3583), rodent CDRs grafted into a human supporting FR prior to fusion with an appropriate human antibody constant domain (Riechmann et al. (1988) *Nature* 332:323-327; Verhoeven et al. (1988) *Science* 239:1534-1536; and Jones et al. (1986) *Nature* 321:522-525), and rodent CDRs supported by recombinantly veneered rodent FRs (European Patent Publication No. 519,596, published Dec. 23, 1992). These "humanized" molecules are designed to minimize unwanted immunological response toward rodent antihuman antibody molecules which limits the duration and effectiveness of therapeutic applications of those moieties in human recipients.

As used herein, the terms "veneered FRs" and "recombinantly veneered FRs" refer to the selective replacement of FR residues from, *e.g.*, a rodent heavy or light chain V region, with human FR residues in order to provide a xenogeneic molecule comprising an antigen-binding site which retains substantially all of the native FR polypeptide folding structure. Veneering techniques are based on the understanding that the ligand binding characteristics of an antigen-binding site are determined primarily by the structure and relative disposition of the heavy and light chain CDR sets within the

antigen-binding surface. Davies et al. (1990) *Ann. Rev. Biochem.* 59:439-473. Thus, antigen binding specificity can be preserved in a humanized antibody only wherein the CDR structures, their interaction with each other, and their interaction with the rest of the V region domains are carefully maintained. By using veneering techniques, exterior  
5 (e.g., solvent-accessible) FR residues which are readily encountered by the immune system are selectively replaced with human residues to provide a hybrid molecule that comprises either a weakly immunogenic, or substantially non-immunogenic veneered surface.

The process of veneering makes use of the available sequence data for  
10 human antibody variable domains compiled by Kabat et al., in *Sequences of Proteins of Immunological Interest*, 4th ed., (U.S. Dept. of Health and Human Services, U.S. Government Printing Office, 1987), updates to the Kabat database, and other accessible U.S. and foreign databases (both nucleic acid and protein). Solvent accessibilities of V region amino acids can be deduced from the known three-dimensional structure for  
15 human and murine antibody fragments. There are two general steps in veneering a murine antigen-binding site. Initially, the FRs of the variable domains of an antibody molecule of interest are compared with corresponding FR sequences of human variable domains obtained from the above-identified sources. The most homologous human V regions are then compared residue by residue to corresponding murine amino acids. The  
20 residues in the murine FR which differ from the human counterpart are replaced by the residues present in the human moiety using recombinant techniques well known in the art. Residue switching is only carried out with moieties which are at least partially exposed (solvent accessible), and care is exercised in the replacement of amino acid residues which may have a significant effect on the tertiary structure of V region  
25 domains, such as proline, glycine and charged amino acids.

In this manner, the resultant "veneered" murine antigen-binding sites are thus designed to retain the murine CDR residues, the residues substantially adjacent to the CDRs, the residues identified as buried or mostly buried (solvent inaccessible), the residues believed to participate in non-covalent (e.g., electrostatic and hydrophobic)  
30 contacts between heavy and light chain domains, and the residues from conserved

structural regions of the FRs which are believed to influence the "canonical" tertiary structures of the CDR loops. These design criteria are then used to prepare recombinant nucleotide sequences which combine the CDRs of both the heavy and light chain of a murine antigen-binding site into human-appearing FRs that can be used to transfect  
5 mammalian cells for the expression of recombinant human antibodies which exhibit the antigen specificity of the murine antibody molecule.

In another embodiment of the invention, monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives  
10 thereof. Preferred radionuclides include  $^{90}\text{Y}$ ,  $^{123}\text{I}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{186}\text{Re}$ ,  $^{188}\text{Re}$ ,  $^{211}\text{At}$ , and  $^{212}\text{Bi}$ . Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, *Pseudomonas* exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-  
20 containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker  
25 group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional  
30 or polyfunctional reagents, both homo- and hetero-functional (such as those described in



the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

5           Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction  
10 of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spittler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

15           It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be  
20 coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

          A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides  
25 such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative  
30 radiohalogenated small molecules and their synthesis. A radionuclide chelate may be

formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

## 5 T Cell Compositions

The present invention, in another aspect, provides T cells specific for a tumor polypeptide disclosed herein, or for a variant or derivative thereof. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone  
10 marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

15 T cells may be stimulated with a polypeptide, polynucleotide encoding a polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide of interest. Preferably, a tumor polypeptide or polynucleotide of the invention is present within a delivery vehicle, such  
20 as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a polypeptide of the present invention if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For  
25 example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen et al., *Cancer Res.* 54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For

example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a tumor polypeptide (100 ng/ml - 100 µg/ml, preferably 200 ng/ml - 25 µg/ml) for 3 - 7 days will typically result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN-γ) is indicative of T cell activation (*see* Coligan et al., Current Protocols in Immunology, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a tumor polypeptide, polynucleotide or polypeptide-expressing APC may be CD4<sup>+</sup> and/or CD8<sup>+</sup>. Tumor polypeptide-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, CD4<sup>+</sup> or CD8<sup>+</sup> T cells that proliferate in response to a tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a tumor polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of the tumor polypeptide can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

#### T Cell Receptor Compositions

The T cell receptor (TCR) consists of 2 different, highly variable polypeptide chains, termed the T-cell receptor α and β chains, that are linked by a disulfide bond (Janeway, Travers, Walport. *Immunobiology*. Fourth Ed., 148-159.

Elsevier Science Ltd/Garland Publishing. 1999). The  $\alpha/\beta$  heterodimer complexes with the invariant CD3 chains at the cell membrane. This complex recognizes specific antigenic peptides bound to MHC molecules. The enormous diversity of TCR specificities is generated much like immunoglobulin diversity, through somatic gene rearrangement. The  $\beta$  chain genes contain over 50 variable (V), 2 diversity (D), over 10 joining (J) segments, and 2 constant region segments (C). The  $\alpha$  chain genes contain over 70 V segments, and over 60 J segments but no D segments, as well as one C segment. During T cell development in the thymus, the D to J gene rearrangement of the  $\beta$  chain occurs, followed by the V gene segment rearrangement to the DJ. This functional VDJ $\beta$  exon is transcribed and spliced to join to a C $\beta$ . For the  $\alpha$  chain, a V $\alpha$  gene segment rearranges to a J $\alpha$  gene segment to create the functional exon that is then transcribed and spliced to the C $\alpha$ . Diversity is further increased during the recombination process by the random addition of P and N-nucleotides between the V, D, and J segments of the  $\beta$  chain and between the V and J segments in the  $\alpha$  chain (Janeway, Travers, Walport. *Immunobiology*. Fourth Ed., 98 and 150. Elsevier Science Ltd/Garland Publishing. 1999).

The present invention, in another aspect, provides TCRs specific for a polypeptide disclosed herein, or for a variant or derivative thereof. In accordance with the present invention, polynucleotide and amino acid sequences are provided for the V-J or V-D-J junctional regions or parts thereof for the alpha and beta chains of the T-cell receptor which recognize tumor polypeptides described herein. In general, this aspect of the invention relates to T-cell receptors which recognize or bind tumor polypeptides presented in the context of MHC. In a preferred embodiment the tumor antigens recognized by the T-cell receptors comprise a polypeptide of the present invention. For example, cDNA encoding a TCR specific for a tumor peptide can be isolated from T cells specific for a tumor polypeptide using standard molecular biological and recombinant DNA techniques.

This invention further includes the T-cell receptors or analogs thereof having substantially the same function or activity as the T-cell receptors of this invention which recognize or bind tumor polypeptides. Such receptors include, but are

not limited to, a fragment of the receptor, or a substitution, addition or deletion mutant of a T-cell receptor provided herein. This invention also encompasses polypeptides or peptides that are substantially homologous to the T-cell receptors provided herein or that retain substantially the same activity. The term "analog" includes any protein or polypeptide having an amino acid residue sequence substantially identical to the T-cell receptors provided herein in which one or more residues, preferably no more than 5 residues, more preferably no more than 25 residues have been conservatively substituted with a functionally similar residue and which displays the functional aspects of the T-cell receptor as described herein.

The present invention further provides for suitable mammalian host cells, for example, non-specific T cells, that are transfected with a polynucleotide encoding TCRs specific for a polypeptide described herein, thereby rendering the host cell specific for the polypeptide. The  $\alpha$  and  $\beta$  chains of the TCR may be contained on separate expression vectors or alternatively, on a single expression vector that also contains an internal ribosome entry site (IRES) for cap-independent translation of the gene downstream of the IRES. Said host cells expressing TCRs specific for the polypeptide may be used, for example, for adoptive immunotherapy of lung cancer as discussed further below.

In further aspects of the present invention, cloned TCRs specific for a polypeptide recited herein may be used in a kit for the diagnosis of lung cancer. For example, the nucleic acid sequence or portions thereof, of tumor-specific TCRs can be used as probes or primers for the detection of expression of the rearranged genes encoding the specific TCR in a biological sample. Therefore, the present invention further provides for an assay for detecting messenger RNA or DNA encoding the TCR specific for a polypeptide.

#### Pharmaceutical Compositions

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions

disclosed herein in pharmaceutically-acceptable carriers for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will be understood that, if desired, a composition as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other  
5 proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from  
10 host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA compositions.

Therefore, in another aspect of the present invention, pharmaceutical compositions are provided comprising one or more of the polynucleotide, polypeptide,  
15 antibody, and/or T-cell compositions described herein in combination with a physiologically acceptable carrier. In certain preferred embodiments, the pharmaceutical compositions of the invention comprise immunogenic polynucleotide and/or polypeptide compositions of the invention for use in prophylactic and therapeutic vaccine applications. Vaccine preparation is generally described in, for example, M.F.  
20 Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Generally, such compositions will comprise one or more polynucleotide and/or polypeptide compositions of the present invention in combination with one or more immunostimulants.

It will be apparent that any of the pharmaceutical compositions described  
25 herein can contain pharmaceutically acceptable salts of the polynucleotides and polypeptides of the invention. Such salts can be prepared, for example, from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.*, salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (*e.g.*, sodium, potassium, lithium, ammonium, calcium and magnesium salts).

In another embodiment, illustrative immunogenic compositions, *e.g.*, vaccine compositions, of the present invention comprise DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the polynucleotide may be administered within any of a variety of delivery systems known to those of ordinary skill in the art. Indeed, numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, 1998, and references cited therein. Appropriate polynucleotide expression systems will, of course, contain the necessary regulatory DNA regulatory sequences for expression in a patient (such as a suitable promoter and terminating signal). Alternatively, bacterial delivery systems may involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope.

Therefore, in certain embodiments, polynucleotides encoding immunogenic polypeptides described herein are introduced into suitable mammalian host cells for expression using any of a number of known viral-based systems. In one illustrative embodiment, retroviruses provide a convenient and effective platform for gene delivery systems. A selected nucleotide sequence encoding a polypeptide of the present invention can be inserted into a vector and packaged in retroviral particles using techniques known in the art. The recombinant virus can then be isolated and delivered to a subject. A number of illustrative retroviral systems have been described (*e.g.*, U.S. Pat. No. 5,219,740; Miller and Rosman (1989) *BioTechniques* 7:980-990; Miller, A. D. (1990) *Human Gene Therapy* 1:5-14; Scarpa et al. (1991) *Virology* 180:849-852; Burns et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:8033-8037; and Boris-Lawrie and Temin (1993) *Cur. Opin. Genet. Develop.* 3:102-109.

In addition, a number of illustrative adenovirus-based systems have also been described. Unlike retroviruses which integrate into the host genome, adenoviruses persist extrachromosomally thus minimizing the risks associated with insertional mutagenesis (Haj-Ahmad and Graham (1986) *J. Virol.* 57:267-274; Bett et al. (1993) *J. Virol.* 67:5911-5921; Mittereder et al. (1994) *Human Gene Therapy* 5:717-729; Seth et al. (1994) *J. Virol.* 68:933-940; Barr et al. (1994) *Gene Therapy* 1:51-58; Berkner, K. L.

(1988) BioTechniques 6:616-629; and Rich et al. (1993) Human Gene Therapy 4:461-476).

Various adeno-associated virus (AAV) vector systems have also been developed for polynucleotide delivery. AAV vectors can be readily constructed using techniques well known in the art. See, e.g., U.S. Pat. Nos. 5,173,414 and 5,139,941; International Publication Nos. WO 92/01070 and WO 93/03769; Lebkowski et al. (1988) Molec. Cell. Biol. 8:3988-3996; Vincent et al. (1990) Vaccines 90 (Cold Spring Harbor Laboratory Press); Carter, B. J. (1992) Current Opinion in Biotechnology 3:533-539; Muzyczka, N. (1992) Current Topics in Microbiol. and Immunol. 158:97-129; Kotin, R. M. (1994) Human Gene Therapy 5:793-801; Shelling and Smith (1994) Gene Therapy 1:165-169; and Zhou et al. (1994) J. Exp. Med. 179:1867-1875.

Additional viral vectors useful for delivering the polynucleotides encoding polypeptides of the present invention by gene transfer include those derived from the pox family of viruses, such as vaccinia virus and avian poxvirus. By way of example, vaccinia virus recombinants expressing the novel molecules can be constructed as follows. The DNA encoding a polypeptide is first inserted into an appropriate vector so that it is adjacent to a vaccinia promoter and flanking vaccinia DNA sequences, such as the sequence encoding thymidine kinase (TK). This vector is then used to transfect cells which are simultaneously infected with vaccinia. Homologous recombination serves to insert the vaccinia promoter plus the gene encoding the polypeptide of interest into the viral genome. The resulting TK<sup>sup</sup>(-) recombinant can be selected by culturing the cells in the presence of 5-bromodeoxyuridine and picking viral plaques resistant thereto.

A vaccinia-based infection/transfection system can be conveniently used to provide for inducible, transient expression or coexpression of one or more polypeptides described herein in host cells of an organism. In this particular system, cells are first infected in vitro with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the polynucleotide or polynucleotides of interest, driven by a T7



promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into polypeptide by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation products. See, *e.g.*, Elroy-Stein and Moss, *Proc. Natl. Acad. Sci. USA* (1990) 87:6743-6747; Fuerst et al. *Proc. Natl. Acad. Sci. USA* (1986) 83:8122-8126.

Alternatively, avipoxviruses, such as the fowlpox and canarypox viruses, can also be used to deliver the coding sequences of interest. Recombinant avipox viruses, expressing immunogens from mammalian pathogens, are known to confer protective immunity when administered to non-avian species. The use of an Avipox vector is particularly desirable in human and other mammalian species since members of the Avipox genus can only productively replicate in susceptible avian species and therefore are not infective in mammalian cells. Methods for producing recombinant Avipoxviruses are known in the art and employ genetic recombination, as described above with respect to the production of vaccinia viruses. See, *e.g.*, WO 91/12882; WO 89/03429; and WO 92/03545.

Any of a number of alphavirus vectors can also be used for delivery of polynucleotide compositions of the present invention, such as those vectors described in U.S. Patent Nos. 5,843,723; 6,015,686; 6,008,035 and 6,015,694. Certain vectors based on Venezuelan Equine Encephalitis (VEE) can also be used, illustrative examples of which can be found in U.S. Patent Nos. 5,505,947 and 5,643,576.

Moreover, molecular conjugate vectors, such as the adenovirus chimeric vectors described in Michael et al. *J. Biol. Chem.* (1993) 268:6866-6869 and Wagner et al. *Proc. Natl. Acad. Sci. USA* (1992) 89:6099-6103, can also be used for gene delivery under the invention.

Additional illustrative information on these and other known viral-based delivery systems can be found, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner et al., *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242;

WO 91/02805; Berkner, *Biotechniques* 6:616-627, 1988; Rosenfeld et al., *Science* 252:431-434, 1991; Kolls et al., *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman et al., *Circulation* 88:2838-2848, 1993; and Guzman et al., *Cir. Res.* 73:1202-1207, 1993.

5           In certain embodiments, a polynucleotide may be integrated into the genome of a target cell. This integration may be in the specific location and orientation via homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the polynucleotide may be stably maintained in the cell as a separate, episomal segment of  
10 DNA. Such polynucleotide segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. The manner in which the expression construct is delivered to a cell and where in the cell the polynucleotide remains is dependent on the type of expression construct employed.

15           In another embodiment of the invention, a polynucleotide is administered/delivered as "naked" DNA, for example as described in Ulmer et al., *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells.

20           In still another embodiment, a composition of the present invention can be delivered via a particle bombardment approach, many of which have been described. In one illustrative example, gas-driven particle acceleration can be achieved with devices such as those manufactured by Powderject Pharmaceuticals PLC (Oxford, UK) and Powderject Vaccines Inc. (Madison, WI), some examples of which are described in  
25 U.S. Patent Nos. 5,846,796; 6,010,478; 5,865,796; 5,584,807; and EP Patent No. 0500 799. This approach offers a needle-free delivery approach wherein a dry powder formulation of microscopic particles, such as polynucleotide or polypeptide particles, are accelerated to high speed within a helium gas jet generated by a hand held device, propelling the particles into a target tissue of interest.

In a related embodiment, other devices and methods that may be useful for gas-driven needle-less injection of compositions of the present invention include those provided by Bioject, Inc. (Portland, OR), some examples of which are described in U.S. Patent Nos. 4,790,824; 5,064,413; 5,312,335; 5,383,851; 5,399,163; 5,520,639  
5 and 5,993,412.

According to another embodiment, the pharmaceutical compositions described herein will comprise one or more immunostimulants in addition to the immunogenic polynucleotide, polypeptide, antibody, T-cell and/or APC compositions of this invention. An immunostimulant refers to essentially any substance that enhances  
10 or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. One preferred type of immunostimulant comprises an adjuvant. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins.  
15 Certain adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated  
20 sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Within certain embodiments of the invention, the adjuvant composition  
25 is preferably one that induces an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (e.g., IFN- $\gamma$ , TNF $\alpha$ , IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (e.g., IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as  
30 provided herein, a patient will support an immune response that includes Th1- and Th2-

type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, 5 *Ann. Rev. Immunol.* 7:145-173, 1989.

Certain preferred adjuvants for eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A, together with an aluminum salt. MPL® adjuvants are available from Corixa Corporation (Seattle, WA; *see*, for example, US 10 Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by 15 Sato et al., *Science* 273:352, 1996. Another preferred adjuvant comprises a saponin, such as Quil A, or derivatives thereof, including QS21 and QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA); Escin; Digitonin; or *Gypsophila* or *Chenopodium quinoa* saponins. Other preferred formulations include more than one saponin in the adjuvant combinations of the present invention, for example 20 combinations of at least two of the following group comprising QS21, QS7, Quil A,  $\beta$ -escin, or digitonin.

Alternatively the saponin formulations may be combined with vaccine vehicles composed of chitosan or other polycationic polymers, polylactide and polylactide-co-glycolide particles, poly-N-acetyl glucosamine-based polymer matrix, 25 particles composed of polysaccharides or chemically modified polysaccharides, liposomes and lipid-based particles, particles composed of glycerol monoesters, etc. The saponins may also be formulated in the presence of cholesterol to form particulate structures such as liposomes or ISCOMs. Furthermore, the saponins may be formulated together with a polyoxyethylene ether or ester, in either a non-particulate solution or 30 suspension, or in a particulate structure such as a paucilamellar liposome or ISCOM. The

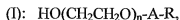
saponins may also be formulated with excipients such as Carbopol<sup>®</sup> to increase viscosity, or may be formulated in a dry powder form with a powder excipient such as lactose.

In one preferred embodiment, the adjuvant system includes the  
5 combination of a monophosphoryl lipid A and a saponin derivative, such as the combination of QS21 and 3D-MPL<sup>®</sup> adjuvant, as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. Another particularly preferred adjuvant formulation employing QS21, 3D-MPL<sup>®</sup> adjuvant and tocopherol in an oil-in-water emulsion is described in WO  
10 95/17210.

Another enhanced adjuvant system involves the combination of a CpG-containing oligonucleotide and a saponin derivative particularly the combination of CpG and QS21 is disclosed in WO 00/09159. Preferably the formulation additionally  
15 comprises an oil in water emulsion and tocopherol.

Additional illustrative adjuvants for use in the pharmaceutical compositions of the invention include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (e.g., SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart,  
20 Belgium), Detox (Enhazyn<sup>®</sup>) (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties, and polyoxyethylene ether adjuvants such as those described in WO 99/52549A1.

25 Other preferred adjuvants include adjuvant molecules of the general formula



wherein,  $n$  is 1-50, A is a bond or  $-\text{C}(\text{O})-$ , R is  $\text{C}_{1-50}$  alkyl or Phenyl  $\text{C}_{1-50}$  alkyl.

One embodiment of the present invention consists of a vaccine  
30 formulation comprising a polyoxyethylene ether of general formula (I), wherein  $n$  is

between 1 and 50, preferably 4-24, most preferably 9; the *R* component is C<sub>1-50</sub>, preferably C<sub>4</sub>-C<sub>20</sub> alkyl and most preferably C<sub>12</sub> alkyl, and *A* is a bond. The concentration of the polyoxyethylene ethers should be in the range 0.1-20%, preferably from 0.1-10%, and most preferably in the range 0.1-1%. Preferred polyoxyethylene  
5 ethers are selected from the following group: polyoxyethylene-9-lauryl ether, polyoxyethylene-9-stearyl ether, polyoxyethylene-8-stearyl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether. Polyoxyethylene ethers such as polyoxyethylene lauryl ether are described in the Merck index (12<sup>th</sup> edition: entry 7717). These adjuvant molecules are described in WO  
10 99/52549.

The polyoxyethylene ether according to the general formula (I) above may, if desired, be combined with another adjuvant. For example, a preferred adjuvant combination is preferably with CpG as described in the pending UK patent application GB 9820956.2.

15 According to another embodiment of this invention, an immunogenic composition described herein is delivered to a host via antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or  
20 maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

25 Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In  
30 general, dendritic cells may be identified based on their typical shape (stellate *in situ*,

with marked cytoplasmic processes (dendrites) visible *in vitro*, their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (*see Zitvogel et al., Nature Med. 4:594-600, 1998*).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNF $\alpha$  to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF $\alpha$ , CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fc $\gamma$  receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (*e.g.*, CD54 and CD11) and costimulatory molecules (*e.g.*, CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide of the invention (or portion or other variant thereof) such that the encoded polypeptide, or an

immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a pharmaceutical composition comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.*, vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (*e.g.*, a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will typically vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, mucosal, intravenous, intracranial, intraperitoneal, subcutaneous and intramuscular administration.

Carriers for use within such pharmaceutical compositions are biocompatible, and may also be biodegradable. In certain embodiments, the formulation preferably provides a relatively constant level of active component release. In other embodiments, however, a more rapid rate of release immediately upon administration may be desired. The formulation of such compositions is well within the level of ordinary skill in the art using known techniques. Illustrative carriers useful in this regard include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other illustrative delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.*,



a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (*see e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

In another illustrative embodiment, biodegradable microspheres (*e.g.*, polylactate polyglycolate) are employed as carriers for the compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344, 5,407,609 and 5,942,252. Modified hepatitis B core protein carrier systems, such as described in WO/99 40934, and references cited therein, will also be useful for many applications. Another illustrative carrier/delivery system employs a carrier comprising particulate-protein complexes, such as those described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

The pharmaceutical compositions of the invention will often further comprise one or more buffers (*e.g.*, neutral buffered saline or phosphate buffered saline), carbohydrates (*e.g.*, glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (*e.g.*, aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate.

The pharmaceutical compositions described herein may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are typically sealed in such a way to preserve the sterility and stability of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a pharmaceutical composition

may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

The development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including  
5 *e.g.*, oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation, is well known in the art, some of which are briefly discussed below for general purposes of illustration.

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions  
10 may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs,  
15 suspensions, syrups, wafers, and the like (see, for example, Mathiowitz *et al.*, Nature 1997 Mar 27;386(6623):410-4; Hwang *et al.*, Crit Rev Ther Drug Carrier Syst 1998;15(3):243-84; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451). Tablets, troches, pills, capsules and the like may also contain any of a variety of additional components, for example, a binder, such as gum tragacanth, acacia,  
20 cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to  
25 materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds  
30 may be incorporated into sustained-release preparation and formulations.

Typically, these formulations will contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of  
5 active compound(s) in each therapeutically useful composition may be prepared in such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a  
10 variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. Alternatively, the active ingredient may be incorporated into an oral solution such as  
15 one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

20 In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even intraperitoneally. Such approaches are well known to the skilled artisan, some of which are further described, for example, in U. S. Patent 5,543,158; U. S. Patent 5,641,515 and U. S. Patent 5,399,363. In certain embodiments, solutions of the active compounds  
25 as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations generally will contain a preservative to prevent the growth of microorganisms.

Illustrative pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (for example, see U. S. Patent 5,466,468). In all cases the form must be sterile and must be fluid to the extent that  
5 easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (e.g., glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable  
10 oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and/or by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be  
15 preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

In one embodiment, for parenteral administration in an aqueous solution,  
20 the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one  
25 dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. Moreover, for human administration, preparations will of

course preferably meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

In another embodiment of the invention, the compositions disclosed herein may be formulated in a neutral or salt form. Illustrative  
5 pharmaceutically-acceptable salts include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium,  
10 calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective.

The carriers can further comprise any and all solvents, dispersion media,  
15 vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary  
20 active ingredients can also be incorporated into the compositions. The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human.

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles.  
25 Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs *via* nasal aerosol sprays has been described, *e.g.*, in U. S. Patent 5,756,353 and U. S. Patent 5,804,212. Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, J Controlled Release 1998 Mar 2;52(1-2):81-7) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871) are also well-known in

the pharmaceutical arts. Likewise, illustrative transmucosal drug delivery in the form of a polytetrafluoroethylene support matrix is described in U. S. Patent 5,780,045.

In certain embodiments, liposomes, nanocapsules, microparticles, lipid particles, vesicles, and the like, are used for the introduction of the compositions of the present invention into suitable host cells/organisms. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like. Alternatively, compositions of the present invention can be bound, either covalently or non-covalently, to the surface of such carrier vehicles.

The formation and use of liposome and liposome-like preparations as potential drug carriers is generally known to those of skill in the art (see for example, Lasic, Trends Biotechnol 1998 Jul;16(7):307-21; Takakura, Nippon Rinsho 1998 Mar;56(3):691-5; Chandran *et al.*, Indian J Exp Biol. 1997 Aug;35(8):801-9; Margalit, Crit Rev Ther Drug Carrier Syst. 1995;12(2-3):233-61; U.S. Patent 5,567,434; U.S. Patent 5,552,157; U.S. Patent 5,565,213; U.S. Patent 5,738,868 and U.S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally difficult to transfect by other procedures, including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, J Biol Chem. 1990 Sep 25;265(27):16337-42; Muller *et al.*, DNA Cell Biol. 1990 Apr;9(3):221-9). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, various drugs, radiotherapeutic agents, enzymes, viruses, transcription factors, allosteric effectors and the like, into a variety of cultured cell lines and animals. Furthermore, the use of liposomes does not appear to be associated with autoimmune responses or unacceptable toxicity after systemic delivery.

In certain embodiments, liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs)).

Alternatively, in other embodiments, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (see, for example, Quintanar-Guerrero *et al.*, Drug Dev Ind Pharm. 5 1998 Dec;24(12):1113-28). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1  $\mu\text{m}$ ) may be designed using polymers able to be degraded *in vivo*. Such particles can be made as described, for example, by Couvreur *et al.*, Crit Rev Ther Drug Carrier Syst. 1988;5(1):1-20; zur Muhlen *et al.*, Eur J Pharm Biopharm. 1998 Mar;45(2):149-55; Zambaux *et al.* J 10 Controlled Release. 1998 Jan 2;50(1-3):31-40; and U. S. Patent 5,145,684.

#### Cancer Therapeutic Methods

Immunologic approaches to cancer therapy are based on the recognition that cancer cells can often evade the body's defenses against aberrant or foreign cells and molecules, and that these defenses might be therapeutically stimulated to regain the 15 lost ground, *e.g.* pgs. 623-648 in Klein, Immunology (Wiley-Interscience, New York, 1982). Numerous recent observations that various immune effectors can directly or indirectly inhibit growth of tumors has led to renewed interest in this approach to cancer therapy, *e.g.* Jager, *et al.*, Oncology 2001;60(1):1-7; Renner, *et al.*, Ann Hematol 2000 Dec;79(12):651-9.

20 Four-basic cell types whose function has been associated with antitumor cell immunity and the elimination of tumor cells from the body are: i) B-lymphocytes which secrete immunoglobulins into the blood plasma for identifying and labeling the nonself invader cells; ii) monocytes which secrete the complement proteins that are responsible for lysing and processing the immunoglobulin-coated target invader cells; 25 iii) natural killer lymphocytes having two mechanisms for the destruction of tumor cells, antibody-dependent cellular cytotoxicity and natural killing; and iv) T-lymphocytes possessing antigen-specific receptors and having the capacity to recognize a tumor cell carrying complementary marker molecules (Schreiber, H., 1989, in Fundamental Immunology (ed). W. E. Paul, pp. 923-955).

Cancer immunotherapy generally focuses on inducing humoral immune responses, cellular immune responses, or both. Moreover, it is well established that induction of CD4<sup>+</sup> T helper cells is necessary in order to secondarily induce either antibodies or cytotoxic CD8<sup>+</sup> T cells. Polypeptide antigens that are selective or ideally  
5 specific for cancer cells, particularly lung cancer cells, offer a powerful approach for inducing immune responses against lung cancer, and are an important aspect of the present invention.

Therefore, in further aspects of the present invention, the pharmaceutical compositions described herein may be used for the treatment of cancer, particularly for  
10 the immunotherapy of lung cancer. Within such methods, the pharmaceutical compositions described herein are administered to a patient, typically a warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. Pharmaceutical  
15 compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. As discussed above, administration of the pharmaceutical compositions may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal,  
20 anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided  
25 herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host  
30 immune system. Examples of effector cells include T cells as discussed above, T



lymphocytes (such as CD8<sup>+</sup> cytotoxic T lymphocytes and CD4<sup>+</sup> T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody  
5 receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

10 Monoclonal antibodies may be labeled with any of a variety of labels for desired selective usages in detection, diagnostic assays or therapeutic applications (as described in U.S. Patent Nos. 6,090,365; 6,015,542; 5,843,398; 5,595,721; and 4,708,930, hereby incorporated by reference in their entirety as if each was incorporated individually). In each case, the binding of the labelled monoclonal antibody to the  
15 determinant site of the antigen will signal detection or delivery of a particular therapeutic agent to the antigenic determinant on the non-normal cell. A further object of this invention is to provide the specific monoclonal antibody suitably labelled for achieving such desired selective usages thereof.

Effector cells may generally be obtained in sufficient quantities for  
20 adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above,  
25 immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known  
30 in the art. For example, antigen-presenting cells can be transfected with a

polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow *in vivo* and to survive  
5 long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see, for example, Cheever et al., Immunological Reviews 157:177, 1997*).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced  
10 into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical  
15 compositions and vaccines may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for  
20 individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor  
25 cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose

ranges from about 25  $\mu$ g to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (e.g., more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

#### Cancer Detection and Diagnostic Compositions, Methods and Kits

In general, a cancer may be detected in a patient based on the presence of one or more lung tumor proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as lung cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample.

Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a tumor sequence should be present at a level that is at least two-fold, preferably three-fold, and more preferably five-fold or higher in tumor tissue than in normal tissue of the same type from which the tumor arose. Expression levels of a particular tumor sequence in tissue types different from that in which the tumor arose are irrelevant in certain diagnostic embodiments since the presence of tumor cells can be confirmed by observation of predetermined differential expression levels, e.g., 2-fold, 5-fold, etc, in tumor tissue to expression levels in normal tissue of the same type.

Other differential expression patterns can be utilized advantageously for diagnostic purposes. For example, in one aspect of the invention, overexpression of a tumor sequence in tumor tissue and normal tissue of the same type, but not in other normal tissue types, e.g. PBMCs, can be exploited diagnostically. In this case, the presence of metastatic tumor cells, for example in a sample taken from the circulation or some other tissue site different from that in which the tumor arose, can be identified and/or confirmed by detecting expression of the tumor sequence in the sample, for example using RT-PCR analysis. In many instances, it will be desired to enrich for tumor cells in the sample of interest, e.g., PBMCs, using cell capture or other like techniques.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. See, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length lung tumor

proteins and polypeptide portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10  $\mu$ g, and preferably about 100 ng to about 1  $\mu$ g, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (*see, e.g.,* Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with lung cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide.

An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For  
5 radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a  
10 specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as lung cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In  
15 one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a  
20 Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off  
25 value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In

general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 $\mu$ g, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use tumor polypeptides to detect antibodies that bind to such polypeptides in a biological



sample. The detection of such tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a tumor protein in a biological sample. Within certain methods, a biological sample comprising CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient is incubated with a tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (e.g., 5 - 25 µg/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of tumor polypeptide to serve as a control. For CD4<sup>+</sup> T cells, activation is preferably detected by evaluating proliferation of the T cells. For CD8<sup>+</sup> T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the tumor protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis.

Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a tumor protein of the invention that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length.

10 In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence as disclosed herein. Techniques for both PCR based assays and hybridization assays are well known in the art (*see*, for example, Mullis et al., *Cold Spring Harbor Symp. Quant. Biol.*, 51:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which

20 may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as

25 compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another aspect of the present invention, cell capture technologies may be used in conjunction, with, for example, real-time PCR to provide a more sensitive tool for detection of metastatic cells expressing lung tumor antigens. Detection of lung

30 cancer cells in biological samples, e.g., bone marrow samples, peripheral blood, and

small needle aspiration samples is desirable for diagnosis and prognosis in lung cancer patients.

Immunomagnetic beads coated with specific monoclonal antibodies to surface cell markers, or tetrameric antibody complexes, may be used to first enrich or positively select cancer cells in a sample. Various commercially available kits may be used, including Dynabeads® Epithelial Enrich (DynaL Biotech, Oslo, Norway), StemSep™ (StemCell Technologies, Inc., Vancouver, BC), and RosetteSep (StemCell Technologies). A skilled artisan will recognize that other methodologies and kits may also be used to enrich or positively select desired cell populations. Dynabeads® Epithelial Enrich contains magnetic beads coated with mAbs specific for two glycoprotein membrane antigens expressed on normal and neoplastic epithelial tissues. The coated beads may be added to a sample and the sample then applied to a magnet, thereby capturing the cells bound to the beads. The unwanted cells are washed away and the magnetically isolated cells eluted from the beads and used in further analyses.

RosetteSep can be used to enrich cells directly from a blood sample and consists of a cocktail of tetrameric antibodies that targets a variety of unwanted cells and crosslinks them to glycophorin A on red blood cells (RBC) present in the sample, forming rosettes. When centrifuged over Ficoll, targeted cells pellet along with the free RBC. The combination of antibodies in the depletion cocktail determines which cells will be removed and consequently which cells will be recovered. Antibodies that are available include, but are not limited to: CD2, CD3, CD4, CD5, CD8, CD10, CD11b, CD14, CD15, CD16, CD19, CD20, CD24, CD25, CD29, CD33, CD34, CD36, CD38, CD41, CD45, CD45RA, CD45RO, CD56, CD66B, CD66e, HLA-DR, IgE, and TCRαβ.

Additionally, it is contemplated in the present invention that mAbs specific for lung tumor antigens can be generated and used in a similar manner. For example, mAbs that bind to tumor-specific cell surface antigens may be conjugated to magnetic beads, or formulated in a tetrameric antibody complex, and used to enrich or positively select metastatic lung tumor cells from a sample. Once a sample is enriched or positively selected, cells may be lysed and RNA isolated. RNA may then be subjected to RT-PCR analysis using lung tumor-specific primers in a real-time PCR

assay as described herein. One skilled in the art will recognize that enriched or selected populations of cells may be analyzed by other methods (*e.g. in situ* hybridization or flow cytometry).

In another embodiment, the compositions described herein may be used  
5 as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the  
10 level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound  
15 binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific  
20 for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided herein may be combined with assays for other known tumor antigens.

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a tumor protein.  
30 Such antibodies or fragments may be provided attached to a support material, as

described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a tumor protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a tumor protein.

The following Examples are offered by way of illustration and not by way of limitation.

## EXAMPLES

### EXAMPLE 1

#### ISOLATION AND CHARACTERIZATION OF cDNA SEQUENCES ENCODING 5 LUNG TUMOR POLYPEPTIDES

This example illustrates the isolation of cDNA molecules encoding lung tumor-specific polypeptides from lung tumor cDNA libraries.

#### A. ISOLATION OF cDNA SEQUENCES FROM A LUNG SQUAMOUS CELL CARCINOMA LIBRARY

- 10 A human lung squamous cell carcinoma cDNA expression library was constructed from poly A<sup>+</sup> RNA from a pool of two patient tissues using a Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning kit (BRL Life Technologies, Gaithersburg, MD) following the manufacturer's protocol. Specifically, lung carcinoma tissues were homogenized with polytron (Kinematica, Switzerland) and total RNA was  
15 extracted using Trizol reagent (BRL Life Technologies) as directed by the manufacturer. The poly A<sup>+</sup> RNA was then purified using an oligo dT cellulose column as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989. First-strand cDNA was synthesized using  
20 the NotI/Oligo-dT18 primer. Double-stranded cDNA was synthesized, ligated with BstXI/EcoRI adaptors (Invitrogen, San Diego, CA) and digested with NotI. Following size fractionation with cDNA size fractionation columns (BRL Life Technologies), the cDNA was ligated into the BstXI/NotI site of pcDNA3.1 (Invitrogen) and transformed into ElectroMax *E. coli* DH10B cells (BRL Life Technologies) by electroporation.

- 25 Using the same procedure, a normal human lung cDNA expression library was prepared from a pool of four tissue specimens. The cDNA libraries were characterized by determining the number of independent colonies, the percentage of clones that carried insert, the average insert size and by sequence analysis. The lung squamous cell carcinoma library contained  $2.7 \times 10^6$  independent colonies, with 100% of clones having an insert and the average insert size being 2100 base pairs. The normal

lung cDNA library contained  $1.4 \times 10^6$  independent colonies, with 90% of clones having inserts and the average insert size being 1800 base pairs. For both libraries, sequence analysis showed that the majority of clones had a full length cDNA sequence and were synthesized from mRNA

- 5 cDNA library subtraction was performed using the above lung squamous cell carcinoma and normal lung cDNA libraries, as described by Hara *et al.* (*Blood*, 84:189-199, 1994) with some modifications. Specifically, a lung squamous cell carcinoma-specific subtracted cDNA library was generated as follows. Normal tissue cDNA library (80  $\mu$ g) was digested with BamHI and XhoI, followed by a filling-in  
10 reaction with DNA polymerase Klenow fragment. After phenol-chloroform extraction and ethanol precipitation, the DNA was dissolved in 133  $\mu$ l of H<sub>2</sub>O, heat-denatured and mixed with 133  $\mu$ l (133  $\mu$ g) of Photoprobe biotin (Vector Laboratories, Burlingame, CA). As recommended by the manufacturer, the resulting mixture was irradiated with a 270 W sunlamp on ice for 20 minutes. Additional Photoprobe biotin (67  $\mu$ l) was added  
15 and the biotinylation reaction was repeated. After extraction with butanol five times, the DNA was ethanol-precipitated and dissolved in 23  $\mu$ l H<sub>2</sub>O to form the driver DNA.

- To form the tracer DNA, 10  $\mu$ g lung squamous cell carcinoma cDNA library was digested with NotI and SpeI, phenol chloroform extracted and passed through Chroma spin-400 columns (Clontech, Palo Alto, CA). Typically, 5  $\mu$ g of  
20 cDNA was recovered after the sizing column. Following ethanol precipitation, the tracer DNA was dissolved in 5  $\mu$ l H<sub>2</sub>O. Tracer DNA was mixed with 15  $\mu$ l driver DNA and 20  $\mu$ l of 2 x hybridization buffer (1.5 M NaCl/10 mM EDTA/50 mM HEPES pH 7.5/0.2% sodium dodecyl sulfate), overlaid with mineral oil, and heat-denatured completely. The sample was immediately transferred into a 68 °C water bath and  
25 incubated for 20 hours (long hybridization [LH]). The reaction mixture was then subjected to a streptavidin treatment followed by phenol/chloroform extraction. This process was repeated three more times. Subtracted DNA was precipitated, dissolved in 12  $\mu$ l H<sub>2</sub>O, mixed with 8  $\mu$ l driver DNA and 20  $\mu$ l of 2 x hybridization buffer, and subjected to a hybridization at 68 °C for 2 hours (short hybridization [SH]). After  
30 removal of biotinylated double-stranded DNA, subtracted cDNA was ligated into

NotI/SpeI site of chloramphenicol resistant pBCSK<sup>+</sup> (Stratagene, La Jolla, CA) and transformed into ElectroMax *E. coli* DH10B cells by electroporation to generate a lung squamous cell carcinoma specific subtracted cDNA library (herein after referred to as "lung subtraction I").

- 5           A second lung squamous cell carcinoma specific subtracted cDNA library (referred to as "lung subtraction II") was generated in a similar way to the lung subtraction library I, except that eight frequently recovered genes from lung subtraction I were included in the driver DNA, and 24,000 independent clones were recovered.

- 10           To analyze the subtracted cDNA libraries, plasmid DNA was prepared from 320 independent clones, randomly picked from the subtracted lung squamous cell carcinoma specific libraries. Representative cDNA clones were further characterized by DNA sequencing with a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A and/or Model 377 (Foster City, CA). The cDNA sequences for sixty isolated clones are provided in SEQ ID NO: 1-60. These sequences were
- 15           compared to known sequences in the gene bank using the EMBL and GenBank databases (release 96). No significant homologies were found to the sequences provided in SEQ ID NO: 2, 3, 19, 38 and 46. The sequences of SEQ ID NO: 1, 6-8, 10-13, 15, 17, 18, 20-27, 29, 30, 32, 34-37, 39-45, 47-49, 51, 52, 54, 55 and 57-59 were found to show some homology to previously identified expressed sequence tags (ESTs).
- 20           The sequences of SEQ ID NO: 9, 28, 31 and 33 were found to show some homology to previously identified non-human gene sequences and the sequences of SEQ ID NO: 4, 5, 14, 50, 53, 56 and 60 were found to show some homology to gene sequences previously identified in humans.

- The subtraction procedure described above was repeated using the above
- 25           lung squamous cell carcinoma cDNA library as the tracer DNA, and the above normal lung tissue cDNA library and a cDNA library from normal liver and heart (constructed from a pool of one sample of each tissue as described above), plus twenty other cDNA clones that were frequently recovered in lung subtractions I and II, as the driver DNA (lung subtraction III). The normal liver and heart cDNA library contained  $1.76 \times 10^6$
- 30           independent colonies, with 100% of clones having inserts and the average insert size



being 1600 base pairs. Ten additional clones were isolated (SEQ ID NO: 61-70). Comparison of these cDNA sequences with those in the gene bank as described above, revealed no significant homologies to the sequences provided in SEQ ID NO: 62 and 67. The sequences of SEQ ID NO: 61, 63-66, 68 and 69 were found to show some  
5 homology to previously isolated ESTs and the sequence provided in SEQ ID NO: 70 was found to show some homology to a previously identified rat gene.

In further studies, the subtraction procedure described above was repeated using the above lung squamous cell carcinoma cDNA library as the tracer DNA, and a cDNA library from a pool of normal lung, kidney, colon, pancreas, brain,  
10 resting PBMC, heart, skin and esophagus as the driver DNA, with esophagus cDNAs making up one third of the driver material. Since esophagus is enriched in normal epithelial cells, including differentiated squamous cells, this procedure is likely to enrich genes that are tumor specific rather than tissues specific. The cDNA sequences of 48 clones determined in this subtraction are provided in SEQ ID NO: 177-224. The  
15 sequences of SEQ ID NO: 177, 178, 180, 181, 183, 187, 192, 195-197, 208, 211, 212, 215, 216, 218 and 219 showed some homology to previously identified genes. The sequences of SEQ ID NO: 179, 182, 184-186, 188-191, 193, 194, 198-207, 209 210, 213, 214, 217, 220 and 224 showed some homology to previously determined ESTs. The sequence of SEQ ID NO: 221-223 showed no homology to any previously  
20 determined sequence.

## B. ISOLATION OF cDNA SEQUENCES FROM A LUNG ADENOCARCINOMA LIBRARY

A human lung adenocarcinoma cDNA expression library was constructed as described above. The library contained  $3.2 \times 10^6$  independent colonies,  
25 with 100% of clones having an insert and the average insert size being 1500 base pairs. Library subtraction was performed as described above using the normal lung and normal liver and heart cDNA expression libraries described above as the driver DNA. Twenty-six hundred independent clones were recovered.

Initial cDNA sequence analysis from 100 independent clones revealed many ribosomal protein genes. The cDNA sequences for fifteen clones isolated in this subtraction are provided in SEQ ID NO: 71-86. Comparison of these sequences with those in the gene bank as described above revealed no significant homologies to the sequence provided in SEQ ID NO: 84. The sequences of SEQ ID NO: 71, 73, 74, 77, 78 and 80-82 were found to show some homology to previously isolated ESTs, and the sequences of SEQ ID NO: 72, 75, 76, 79, 83 and 85 were found to show some homology to previously identified human genes.

In further studies, a cDNA library (referred to as mets3616A) was constructed from a metastatic lung adenocarcinoma. The determined cDNA sequences of 25 clones sequenced at random from this library are provided in SEQ ID NO: 255-279. The mets3616A cDNA library was subtracted against a cDNA library prepared from a pool of normal lung, liver, pancreas, skin, kidney, brain and resting PBMC. To increase the specificity of the subtraction, the driver was spiked with genes that were determined to be most abundant in the mets3616A cDNA library, such as EF1-alpha, integrin-beta and anticoagulant protein PP4, as well as with cDNAs that were previously found to be differentially expressed in subtracted lung adenocarcinoma cDNA libraries. The determined cDNA sequences of 51 clones isolated from the subtracted library (referred to as mets3616A-S1) are provided in SEQ ID NO: 280-330.

Comparison of the sequences of SEQ ID NO: 255-330 with those in the public databases revealed no significant homologies to the sequences of SEQ ID NO: 255-258, 260, 262-264, 270, 272, 275, 276, 279, 281, 287, 291, 296, 300 and 310. The sequences of SEQ ID NO: 259, 261, 265-269, 271, 273, 274, 277, 278, 282-285, 288-290, 292, 294, 297-299, 301, 303-309, 313, 314, 316, 320-324 and 326-330 showed some homology to previously identified gene sequences, while the sequences of SEQ ID NO: 280, 286, 293, 302, 310, 312, 315, 317-319 and 325 showed some homology to previously isolated expressed sequence tags (ESTs).

## EXAMPLE 2

DETERMINATION OF TISSUE SPECIFICITY OF LUNG TUMOR POLYPEPTIDES

Using gene specific primers, mRNA expression levels for seven representative lung tumor polypeptides described in Example 1 were examined in a  
5 variety of normal and tumor tissues using RT-PCR.

Briefly, total RNA was extracted from a variety of normal and tumor tissues using Trizol reagent as described above. First strand synthesis was carried out using 2 µg of total RNA with SuperScript II reverse transcriptase (BRL Life Technologies) at 42 °C for one hour. The cDNA was then amplified by PCR with gene-  
10 specific primers. To ensure the semi-quantitative nature of the RT-PCR, β-actin was used as an internal control for each of the tissues examined. 1 µl of 1:30 dilution of cDNA was employed to enable the linear range amplification of the β-actin template and was sensitive enough to reflect the differences in the initial copy numbers. Using these conditions, the β-actin levels were determined for each reverse transcription  
15 reaction from each tissue. DNA contamination was minimized by DNase treatment and by assuring a negative PCR result when using first strand cDNA that was prepared without adding reverse transcriptase.

mRNA Expression levels were examined in five different types of tumor tissue (lung squamous cell carcinoma from 3 patients, lung adenocarcinoma, colon  
20 tumor from 2 patients, breast tumor and prostate tumor), and thirteen different normal tissues (lung from 4 donors, prostate, brain, kidney, liver, ovary, skeletal muscle, skin, small intestine, stomach, myocardium, retina and testes). Using a 10-fold amount of cDNA, the antigen LST-S1-90 (SEQ ID NO: 3) was found to be expressed at high levels in lung squamous cell carcinoma and in breast tumor, and at low to undetectable  
25 levels in the other tissues examined.

The antigen LST-S2-68 (SEQ ID NO: 15) appears to be specific to lung and breast tumor, however, expression was also detected in normal kidney. Antigens LST-S1-169 (SEQ ID NO: 6) and LST-S1-133 (SEQ ID NO: 5) appear to be very abundant in lung tissues (both normal and tumor), with the expression of these two  
30 genes being decreased in most of the normal tissues tested. Both LST-S1-169 and LST-

S1-133 were also expressed in breast and colon tumors. Antigens LST-S1-6 (SEQ ID NO: 7) and LST-S2-I2-5F (SEQ ID NO: 47) did not show tumor or tissue specific expression, with the expression of LST-S1-28 being rare and only detectable in a few tissues. The antigen LST-S3-7 (SEQ ID NO: 63) showed lung and breast tumor specific expression, with its message only being detected in normal testes when the PCR was performed for 30 cycles. Lower level expression was detected in some normal tissues when the cycle number was increased to 35. Antigen LST-S3-13 (SEQ ID NO: 66) was found to be expressed in 3 out of 4 lung tumors, one breast tumor and both colon tumor samples. Its expression in normal tissues was lower compared to tumors, and was only detected in 1 out of 4 normal lung tissues and in normal tissues from kidney, ovary and retina. Expression of antigens LST-S3-4 (SEQ ID NO: 62) and LST-S3-14 (SEQ ID NO: 67) was rare and did not show any tissue or tumor specificity. Consistent with Northern blot analyses, the RT-PCR results on antigen LAT-S1-A-10A (SEQ ID NO: 78) suggested that its expression is high in lung, colon, stomach and small intestine tissues, including lung and colon tumors, whereas its expression was low or undetectable in other tissues.

A total of 2002 cDNA fragments isolated in lung subtractions I, II and III, described above, were colony PCR amplified and their mRNA expression levels in lung tumor, normal lung, and various other normal and tumor tissues were determined using microarray technology (Synteni, Palo Alto, CA). Briefly, the PCR amplification products were dotted onto slides in an array format, with each product occupying a unique location in the array. mRNA was extracted from the tissue sample to be tested, reverse transcribed, and fluorescent-labeled cDNA probes were generated. The microarrays were probed with the labeled cDNA probes, the slides scanned and fluorescence intensity was measured. This intensity correlates with the hybridization intensity. Seventeen non-redundant cDNA clones showed over-expression in lung squamous tumors, with expression in normal tissues tested (lung, skin, lymph node, colon, liver, pancreas, breast, heart, bone marrow, large intestine, kidney, stomach, brain, small intestine, bladder and salivary gland) being either undetectable, or 10-fold less compared to lung squamous tumors. The determined cDNA sequences for the

clone L513S are provided in SEQ ID NO: 87 and 88; those for L514S are provided in SEQ ID NO: 89 and 90; those for L516S in SEQ ID NO: 91 and 92; that for L517S in SEQ ID NO: 93; that for L519S in SEQ ID NO: 94; those for L520S in SEQ ID NO: 95 and 96; those for L521S in SEQ ID NO: 97 and 98; that for L522S in SEQ ID NO: 99; 5 that for L523S in SEQ ID NO: 100; that for L524S in SEQ ID NO: 101; that for L525S in SEQ ID NO: 102; that for L526S in SEQ ID NO: 103; that for L527S in SEQ ID NO: 104; that for L528S in SEQ ID NO: 105; that for L529S in SEQ ID NO: 106; and those for L530S in SEQ ID NO: 107 and 108. Additionally, the full-length cDNA sequence for L530S is provided in SEQ ID NO: 151, with the corresponding amino acid sequence 10 being provided in SEQ ID NO: 152. L530S shows homology to a splice variant of a p53 tumor suppressor homologue, p63. The cDNA sequences of 7 known isoforms of p63 are provided in SEQ ID NO: 331-337, with the corresponding amino acid sequences being provided in SEQ ID NO: 338-344, respectively.

Due to polymorphisms, the clone L531S appears to have two forms. A 15 first determined full-length cDNA sequence for L531S is provided in SEQ ID NO: 109, with the corresponding amino acid sequence being provided in SEQ ID NO: 110. A second determined full-length cDNA sequence for L531S is provided in SEQ ID NO: 111, with the corresponding amino acid sequence being provided in SEQ ID NO: 112. The sequence of SEQ ID NO: 111 is identical to that of SEQ ID NO: 109, except that it 20 contains a 27 bp insertion. Similarly, L514S has two alternatively spliced forms; the first variant cDNA is listed as SEQ ID NO: 153, with the corresponding amino acid sequence being provided in SEQ ID NO: 155. The full-length cDNA for the second variant form of L514S is provided in SEQ ID NO: 154, with the corresponding amino acid sequence being provided in SEQ ID NO: 156.

25 Full length cloning for L524S (SEQ ID NO: 101) yielded two variants (SEQ ID NO: 163 and 164) with the corresponding amino acid sequences of SEQ ID NO: 165 and 166, respectively. Both variants have been shown to encode parathyroid hormone-related peptide.

Attempts to isolate the full-length cDNA for L519S, resulted in the 30 isolation of the extended cDNA sequence provided in SEQ ID NO: 173, which contains

a potential open reading frame. The amino acid sequence encoded by the sequence of SEQ ID NO: 173 is provided in SEQ ID NO: 174. Additionally, the full-length cDNA sequence for the clone of SEQ ID NO: 100 (known as L523S), a known gene, is provided in SEQ ID NO: 175, with the corresponding amino acid sequence being  
5 provided in SEQ ID NO: 176. In further studies, a full-length cDNA sequence for L523S was isolated from a L523S-positive tumor cDNA library by PCR amplification using gene specific primers designed from the sequence of SEQ ID NO: 175. The determined full-length cDNA sequence is provided in SEQ ID NO: 347. The amino acid sequence encoded by this sequence is provided in SEQ ID NO: 348. This protein  
10 sequence differs from the previously published protein sequence at two amino acid positions, namely at positions 158 and 410.

Comparison of the sequences of L514S and L531S (SEQ ID NO: 87 and 88, and 109, respectively) with those in the gene bank, as described above, revealed no significant homologies to known sequences. The sequences of L513S, L516S, L517S,  
15 L519S, L520S and L530S (SEQ ID NO: 87 and 88, 91 and 92, 93, 94, 95 and 96, 107 and 108, respectively) were found to show some homology to previously identified ESTs. The sequences of L521S, L522S, L523S, L524S, L525S, L526S, L527S, L528S and L529S (SEQ ID NO: 97 and 98, 99, 99, 101, 102, 103, 104, 105, and 106, respectively) were found to represent known genes. The determined full-length cDNA  
20 sequence for L520S is provided in SEQ ID NO: 113, with the corresponding amino acid sequence being provided in SEQ ID NO: 114. Subsequent microarray analysis showed L520S to be overexpressed in breast tumors in addition to lung squamous tumors.

Further analysis demonstrated that L529S (SEQ ID NO: 106 and 115),  
25 L525S (SEQ ID NO: 102 and 120) and L527S (SEQ ID NO: 104) are cytoskeletal components and potentially squamous cell specific proteins. L529S is connexin 26, a gap junction protein. It was found to be highly expressed in one lung squamous tumor, referred to as 9688T, and moderately over-expressed in two others. However, lower level expression of connexin 26 is also detectable in normal skin, colon, liver and  
30 stomach. The over-expression of connexin 26 in some breast tumors has been reported

and a mutated form of L529S may result in over-expression in lung tumors. L525S is plakophilin 1, a desmosomal protein found in plaque-bearing adhering junctions of the skin. Expression levels for L525S mRNA was highly elevated in three out of four lung squamous tumors tested, and in normal skin. L527S has been identified as keratin 6 isoform, type II 58 Kd keratin and cytokeratin 13, and shows over-expression in squamous tumors and low expression in normal skin, breast and colon tissues. Keratin and keratin-related genes have been extensively documented as potential markers for lung cancer including CYFRA2.1 (Pastor, A., et al, *Eur. Respir. J.*, 10:603-609, 1997). L513S (SEQ ID NO: 87 and 88) shows moderate over-expression in several tumor tissues tested, and encodes a protein that was first isolated as a pemphigus vulgaris antigen.

L520S (SEQ ID NO: 95 and 96) and L521S (SEQ ID NO: 97 and 98) are highly expressed in lung squamous tumors, with L520S being up-regulated in normal salivary gland and L521S being over-expressed in normal skin. Both belong to a family of small proline rich proteins and represent markers for fully differentiated squamous cells. L521S has been described as a specific marker for lung squamous tumor (Hu, R., et al, *Lung Cancer*, 20:25-30, 1998). L515S (SEQ ID NO: 162) encodes IGF- $\beta$ 2 and L516S is an aldose reductase homologue. Both are moderately expressed in lung squamous tumors and in normal colon. Notably, L516S (SEQ ID NO: 91 and 92) is up-regulated in metastatic tumors but not primary lung adenocarcinoma, an indication of its potential role in metatasis and a potential prognostic marker. L522S (SEQ ID NO: 99) is moderately over-expressed in lung squamous tumors with minimum expression in normal tissues. L522S has been shown to belong to a class IV alcohol dehydrogenase, ADH7, and its expression profile suggests it is a squamous cell specific antigen. L523S (SEQ ID NO: 100) is moderately over-expressed in lung squamous tumor, human pancreatic cancer cell lines and pancreatic cancer tissues, suggesting this gene may be a shared antigen between pancreatic and lung squamous cell cancer.

L524S (SEQ ID NO: 101) is over-expressed in the majority of squamous tumors tested and is homologous with parathyroid hormone-related peptide (PTHrP), which is best known to cause humoral hypercalcaemia associated with malignant

tumors such as leukemia, prostate and breast cancer. It is also believed that PTHrP is most commonly associated with squamous carcinoma of lung and rarely with lung adenocarcinoma (Davidson, L.A., et al, *J. Pathol.*, 178: 398-401, 1996). L528S (SEQ ID NO: 105) is highly over-expressed in two lung squamous tumors with moderate expression in two other squamous tumors, one lung adenocarcinoma and some normal tissues, including skin, lymph nodes, heart, stomach and lung. It encodes the NMB gene that is similar to the precursor of melanocyte specific gene Pmel17, which is reported to be preferentially expressed in low-metastatic potential melanoma cell lines. This suggests that L528S may be a shared antigen in both melanoma and lung squamous cell carcinoma. L526S (SEQ ID NO: 103) was overexpressed in all lung squamous cell tumor tissues tested and has been shown to share homology with a gene (ATM) in which a mutation causes ataxia telangiectasia, a genetic disorder in humans causing a predisposition to cancer, among other symptoms. ATM encodes a protein that activates a p53 mediated cell-cycle checkpoint through direct binding and phosphorylation of the p53 molecule. Approximately 40% of lung cancers are associated with p53 mutations, and it is speculated that over-expression of ATM is a result of compensation for loss of p53 function, but it is unknown whether over-expression is the cause of result of lung squamous cell carcinoma. Additionally, expression of L526S (ATM) is also detected in a metastatic but not lung adenocarcinoma, suggesting a role in metastasis.

Expression of L523S (SEQ ID NO: 175), was examined by real time RT-PCR as described above. In a first study using a panel of lung squamous tumors, L523S was found to be expressed in 4/7 lung squamous tumors, 2/3 head and neck squamous tumors and 2/2 lung adenocarcinomas, with low level expression being observed in skeletal muscle, soft palate and tonsil. In a second study using a lung adenocarcinoma panel, expression of L523S was observed in 4/9 primary adenocarcinomas, 2/2 lung pleural effusions, 1/1 metastatic lung adenocarcinomas and 2/2 lung squamous tumors, with little expression being observed in normal tissues.

Expression of L523S in lung tumors and various normal tissues was also examined by Northern blot analysis, using standard techniques. In a first study, L523S was found to be expressed in a number of lung adenocarcinomas and squamous cell



carcinomas, as well as normal tonsil. No expression was observed in normal lung. In a second study using a normal tissue blot (referred to as HB-12) from Clontech, no expression was observed in brain, skeletal muscle, colon, thymus, spleen, kidney, liver, small intestine, lung or PBMC, although there was strong expression in placenta.

5

### EXAMPLE 3

#### ISOLATION AND CHARACTERIZATION OF LUNG TUMOR POLYPEPTIDES BY PCR-BASED SUBTRACTION

Eight hundred and fifty seven clones from a cDNA subtraction library, containing cDNA from a pool of two human lung squamous tumors subtracted against eight normal human tissue cDNAs including lung, PBMC, brain, heart, kidney, liver, pancreas, and skin, (Clontech, Palo Alto, CA) were derived and submitted to a first round of PCR amplification. This library was subjected to a second round of PCR amplification, following the manufacturer's protocol. The resulting cDNA fragments were subcloned into the P7-Adv vector (Clontech, Palo Alto, CA) and transformed into DH5 $\alpha$  *E. coli* (Gibco, BRL). DNA was isolated from independent clones and sequenced using a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A.

One hundred and sixty two positive clones were sequenced. Comparison of the DNA sequences of these clones with those in the EMBL and GenBank databases, as described above, revealed no significant homologies to 13 of these clones, hereinafter referred to as Contigs 13, 16, 17, 19, 22, 24, 29, 47, 49, 56-59. The determined cDNA sequences for these clones are provided in SEQ ID NO: 125, 127-129, 131-133, 142, 144, 148-150, and 157, respectively. Contigs 1, 3-5, 7-10, 12, 11, 15, 20, 31, 33, 38, 39, 41, 43, 44, 45, 48, 50, 53, 54 (SEQ ID NO: 115-124, 126, 130, 134-141, 143, 145-147, respectively) were found to show some degree of homology to previously identified DNA sequences. Contig 57 (SEQ ID NO: 149) was found to represent the clone L519S (SEQ ID NO: 94) disclosed in US. Patent Application No. 09/123,912, filed July 27, 1998. To the best of the inventors' knowledge, none of these sequences have been previously shown to be differentially over-expressed in lung tumors.

mRNA expression levels for representative clones in lung tumor tissues, normal lung tissues (n=4), resting PBMC, salivary gland, heart, stomach, lymph nodes, skeletal muscle, soft palate, small intestine, large intestine, bronchial, bladder, tonsil, kidney, esophagus, bone marrow, colon, adrenal gland, pancreas, and skin (all derived from human) were determined by RT-PCR as described above. Expression levels using microarray technology, as described above, were examined in one sample of each tissue type unless otherwise indicated.

Contig 3 (SEQ ID NO: 116) was found to be highly expressed in all head and neck squamous cell tumors tested (17/17), and expressed in the majority (8/12) of lung squamous tumors, (high expression in 7/12, moderate in 2/12, and low in 2/12), while showing negative expression for 2/4 normal lung tissues and low expression in the remaining two samples. Contig 3 showed moderate expression in skin and soft palate, and lowered expression levels in resting PBMC, large intestine, salivary gland, tonsil, pancreas, esophagus, and colon. Contig 11 (SEQ ID NO: 124) was found to be expressed in all head and neck squamous cell tumors tested (17/17), with high levels of expression being seen in 14/17 tumors, and moderately levels of expression being seen in 3/17 tumors. Additionally, high expression was seen in 3/12 lung squamous tumors and moderate expression in 4/12 lung squamous tumors. Contig 11 was negative for 3/4 normal lung samples, with the remaining sample having only low expression. Contig 11 showed low to moderate reactivity to salivary gland, soft palate, bladder, tonsil, skin, esophagus, and large intestine. Contig 13 (SEQ ID NO: 125) was found to be expressed in all head and neck squamous cell tumors tested (17/17), with high expression in 12/17, and moderate expression in 5/17. Contig 13 was expressed in 7/12 lung squamous tumors, with high expression in 4/12 and moderate expression in three samples. Analysis of normal lung samples showed negative expression for 2/4 and low to moderate expression in the remaining two samples. Contig 13 showed low to moderate reactivity to resting PBMC, salivary gland, bladder, pancreas, tonsil, skin, esophagus, and large intestine, as well as high expression in soft palate. Subsequent full-length cloning efforts revealed that contig 13 (also known as L761P) maps to the 3'

untranslated region of the hSec10p gene. The full-length sequence for this gene is set forth in SEQ ID NO: 368, and encodes the protein set forth in SEQ ID NO: 369.

Contig 16 (SEQ ID NO: 127) was found to be moderately expressed in several head and neck squamous cell tumors (6/17) and one lung squamous tumor, while showing no expression in any normal lung samples tested. Contig 16 showed low reactivity to resting PBMC, large intestine, skin, salivary gland, and soft palate. Contig 17 (SEQ ID NO: 128) was shown to be expressed in all head and neck squamous cell tumors tested (17/17) (highly expressed in 5/17, and moderately expressed in 12/17). Determination of expression levels in lung squamous tumors showed one tumor sample with high expression and 3/12 with moderate levels. Contig 17 was negative for 2/4 normal lung samples, with the remaining samples having only low expression. Additionally, low level expression was found in esophagus and soft palate. Contig 19 (SEQ ID NO: 129) was found to be expressed in most head and neck squamous cell tumors tested (11/17); with two samples having high expression levels, 6/17 showing moderate expression, and low expression being found in 3/17. Testing in lung squamous tumors revealed only moderate expression in 3/12 samples. Expression levels in 2/4 of normal lung samples were negative, the two other samples having only low expression. Contig 19 showed low expression levels in esophagus, resting PBMC, salivary gland, bladder, soft palate and pancreas.

Contig 22 (SEQ ID NO: 131), was shown to be expressed in most head and neck squamous cell tumors tested (13/17) with high expression in four of these samples, moderate expression in 6/17, and low expression in 3/17. Expression levels in lung squamous tumors were found to be moderate to high for 3/12 tissues tested, with negative expression in two normal lung samples and low expression in two other samples (n=4). Contig 22 showed low expression in skin, salivary gland and soft palate. Similarly, Contig 24 (SEQ ID NO: 132) was found to be expressed in most head and neck squamous cell tumors tested (13/17) with high expression in three of these samples, moderate expression in 6/17, and low expression in 4/17. Expression levels in lung squamous tumors were found to be moderate to high for 3/12 tissues tested, with negative expression for three normal lung samples and low expression in one sample

(n=4). Contig 24 showed low expression in skin, salivary gland and soft palate. Contig 29 (SEQ ID NO: 133) was expressed in nearly all head and neck squamous cell tumors tested (16/17): highly expressed in 4/17, moderately expressed in 11/17, with low expression in one sample. Also, it was moderately expressed in 3/12 lung squamous tumors, while being negative for 2/4 normal lung samples. Contig 29 showed low to moderate expression in large intestine, skin, salivary gland, pancreas, tonsil, heart and soft palate. Contig 47 (SEQ ID NO: 142) was expressed in most head and neck squamous cell tumors tested (12/17): moderate expression in 10/17, and low expression in two samples. In lung squamous tumors, it was highly expressed in one sample and moderately expressed in two others (n=13). Contig 47 was negative for 2/4 normal lung samples, with the remaining two samples having moderate expression. Also, Contig 47 showed moderate expression in large intestine, and pancreas, and low expression in skin, salivary gland, soft palate, stomach, bladder, resting PBMC, and tonsil.

Contig 48 (SEQ ID NO: 143) was expressed in all head and neck squamous cell tumors tested (17/17): highly expressed in 8/17 and moderately expressed in 7/17, with low expression in two samples. Expression levels in lung squamous tumors were high to moderate in three samples (n=13). Contig 48 was negative for one out of four normal lung samples, the remaining showing low or moderate expression. Contig 48 showed moderate expression in soft palate, large intestine, pancreas, and bladder, and low expression in esophagus, salivary gland, resting PBMC, and heart. Contig 49 (SEQ ID NO: 144) was expressed at low to moderate levels in 6/17 head and neck squamous cell tumors tested. Expression levels in lung squamous tumors were moderate in three samples (n=13). Contig 49 was negative for 2/4 normal lung samples, the remaining samples showing low expression. Moderate expression levels in skin, salivary gland, large intestine, pancreas, bladder and resting PBMC were shown, as well as low expression in soft palate, lymph nodes, and tonsil. Contig 56 (SEQ ID NO: 148) was expressed in low to moderate levels in 3/17 head and neck squamous cell tumors tested, and in lung squamous tumors, showing low to moderate levels in three out of thirteen samples. Notably, low expression levels were

detected in one adenocarcinoma lung tumor sample (n=2). Contig 56 was negative for 3/4 normal lung samples, and showed moderate expression levels in only large intestine, and low expression in salivary gland, soft palate, pancreas, bladder, and resting PBMC. Contig 58, also known as L769P, (SEQ ID NO: 150) was expressed at moderate levels in 11/17 head and neck squamous cell tumors tested and low expression in one additional sample. Expression in lung squamous tumors showed low to moderate levels in three out of thirteen samples. Contig 58 was negative for 3/4 normal lung samples, with one sample having low expression. Moderate expression levels in skin, large intestine, and resting PBMC were demonstrated, as well as low expression in salivary gland, soft palate, pancreas, and bladder. Contig 59 (SEQ ID NO: 157) was expressed in some head, neck, and lung squamous tumors. Low level expression of Contig 59 was also detected in salivary gland and large intestine.

The full-length cDNA sequence for Contig 22, also referred to as L763P, is provided in SEQ ID NO: 158, with the corresponding amino acid sequence being provided in SEQ ID NO: 159. Real-time RT-PCR analysis of L763P revealed that it is highly expressed in 3/4 lung squamous tumors as well as 4/4 head and neck squamous tumors, with low level expression being observed in normal brain, skin, soft pallet and trachea. Subsequent database searches revealed that the sequence of SEQ ID NO: 158 contains a mutation, resulting in a frameshift in the corresponding protein sequence. A second cDNA sequence for L763P is provided in SEQ ID NO: 345, with the corresponding amino acid sequence being provided in SEQ ID NO: 346. The sequences of SEQ ID NO: 159 and 346 are identical with the exception of the C-terminal 33 amino acids of SEQ ID NO: 159.

The full-length cDNA sequence incorporating Contigs 17, 19, and 24, referred to as L762P, is provided in SEQ ID NO: 160, with the corresponding amino acid sequence being provided in SEQ ID NO: 161. Further analysis of L762P has determined it to be a type I membrane protein and two additional variants have been sequenced. Variant 1 (SEQ ID NO: 167, with the corresponding amino acid sequence in SEQ ID NO: 169) is an alternatively spliced form of SEQ ID NO: 160 resulting in deletion of 503 nucleotides, as well as deletion of a short segment of the expressed

protein. Variant 2 (SEQ ID NO: 168, with the corresponding amino acid sequence in SEQ ID NO: 170) has a two nucleotide deletion at the 3' coding region in comparison to SEQ ID NO: 160, resulting in a secreted form of the expressed protein. Real-time RT-PCR analysis of L762P revealed that is over-expressed in 3/4 lung squamous tumors and 4/4 head & neck tumors, with low level expression being observed in normal skin, soft pallet and trachea.

An epitope of L762P was identified as having the sequence KPGHWTYTLNNTHHSLQALK (SEQ ID NO: 382), which corresponds to amino acids 571-590 of SEQ ID NO:161.

The full-length cDNA sequence for contig 56 (SEQ ID NO: 148), also referred to as L773P, is provided in SEQ ID NO: 171, with the amino acid sequence in SEQ ID NO: 172. L773P was found to be identical to dihydroxyl dehydrogenase at the 3' portion of the gene, with divergent 5' sequence. As a result, the 69 N-terminal amino acids are unique. The cDNA sequence encoding the 69 N-terminal amino acids is provided in SEQ ID NO: 349, with the N-terminal amino acid sequence being provided in SEQ ID NO: 350. Real-time PCR revealed that L773P is highly expressed in lung squamous tumor and lung adenocarcinoma, with no detectable expression in normal tissues. Subsequent Northern blot analysis of L773P demonstrated that this transcript is differentially over-expressed in squamous tumors and detected at approximately 1.6 Kb in primary lung tumor tissue and approximately 1.3 Kb in primary head and neck tumor tissue.

Subsequent microarray analysis has shown Contig 58, also referred to as L769S (SEQ ID NO: 150), to be overexpressed in breast tumors in addition to lung squamous tumors.

#### EXAMPLE 4

##### ISOLATION AND CHARACTERIZATION OF LUNG TUMOR POLYPEPTIDES BY PCR-BASED SUBTRACTION

Seven hundred and sixty clones from a cDNA subtraction library, containing cDNA from a pool of two human lung primary adenocarcinomas subtracted

against a pool of nine normal human tissue cDNAs including skin, colon, lung, esophagus, brain, kidney, spleen, pancreas and liver, (Clontech, Palo Alto, CA) were derived and submitted to a first round of PCR amplification. This library (referred to as ALT-1) was subjected to a second round of PCR amplification, following the manufacturer's protocol. The expression levels of these 760 cDNA clones in lung tumor, normal lung, and various other normal and tumor tissues, were examined using microarray technology (Incyte, Palo Alto, CA). Briefly, the PCR amplification products were dotted onto slides in an array format, with each product occupying a unique location in the array. mRNA was extracted from the tissue sample to be tested, reverse transcribed, and fluorescent-labeled cDNA probes were generated. The microarrays were probed with the labeled cDNA probes, the slides scanned and fluorescence intensity was measured. This intensity correlates with the hybridization intensity.. A total of 118 clones, of which 55 were unique, were found to be over-expressed in lung tumor tissue, with expression in normal tissues tested (lung, skin, lymph node, colon, liver, pancreas, breast, heart, bone marrow, large intestine, kidney, stomach, brain, small intestine, bladder and salivary gland) being either undetectable, or at significantly lower levels. One of these clones, having the sequence as provided in SEQ ID NO:420 (clone #19014), shows homology to a previously identified clone, L773P. Clone L773P has the full-length cDNA sequence provided in SEQ ID NO:171 and the amino acid sequence provided in SEQ ID NO:172 The isolation of clone #19014 is also described in co-pending U.S. Patent application 09/285,479, filed April 2, 1999.

## EXAMPLE 5

### SYNTHESIS OF POLYPEPTIDES

Polypeptides may be synthesized on a Perkin Elmer/Applied Biosystems Division 430A peptide synthesizer using Fmoc chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support is carried out using the following

cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides are precipitated in cold methyl-t-butyl-ether. The peptide pellets are then dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides are characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

10

## EXAMPLE 6

PREPARATION OF ANTIBODIES AGAINST LUNG CANCER ANTIGENS

Polyclonal antibodies against the lung cancer antigens L514S, L528S, L531S, L523 and L773P (SEQ ID NO: 155, 225, 112, 176 and 171, respectively) were prepared as follows.

15

Rabbits were immunized with recombinant protein expressed in and purified from *E. coli* as described below. For the initial immunization, 400 µg of antigen combined with muramyl dipeptide (MDP) was injected subcutaneously (S.C.). Animals were boosted S.C. 4 weeks later with 200 µg of antigen mixed with incomplete Freund's Adjuvant (IFA). Subsequent boosts of 100 µg of antigen mixed with IFA were injected S.C. as necessary to induce high antibody titer responses. Serum bleeds from immunized rabbits were tested for antigen-specific reactivity using ELISA assays with purified protein. Polyclonal antibodies against L514S, L528S, L531S, L523S and L773P were affinity purified from high titer polyclonal sera using purified protein attached to a solid support.

25

Immunohistochemical analysis using polyclonal antibodies against L514S was performed on a panel of 5 lung tumor samples, 5 normal lung tissue samples and normal colon, kidney, liver, brain and bone marrow. Specifically, tissue samples were fixed in formalin solution for 24 hours and embedded in paraffin before being sliced into 10 micron sections. Tissue sections were permeabilized and incubated with antibody for 1 hr. HRP-labeled anti-mouse followed by incubation with DAB

30



chromogen was used to visualize L514S immunoreactivity. L514S was found to be highly expressed in lung tumor tissue with little or no expression being observed in normal lung, brain or bone marrow. Light staining was observed in colon (epithelial crypt cells positive) and kidney (tubules positive). Staining was seen in normal liver but  
5 no mRNA has been detected in this tissue making this result suspect.

Using the same procedure, immunohistochemical analysis using polyclonal antibodies against L528S demonstrated staining in lung tumor and normal lung samples, light staining in colon and kidney, and no staining in liver and heart.

Immunohistochemical analysis using polyclonal antibodies against  
10 L531S demonstrated staining in lung tumor samples, light membrane staining in most normal lung samples, epithelial staining in colon, tubule staining in kidney, ductal epithelial staining in liver and no staining in heart.

Immunohistochemical analysis using polyclonal antibodies against L523S demonstrated staining in all lung cancer samples tested but no staining in normal  
15 lung, kidney, liver, colon, bone marrow or cerebellum.

Generation of polyclonal anti-sera against L762P (SEQ ID NO: 169 and 170) was performed as follows. 400 micrograms of lung antigen was combined with 100 micrograms of muramyl dipeptide (MDP). An equal volume of Incomplete Freund's Adjuvant (IFA) was added and then mixed until an emulsion was formed.  
20 Rabbits were injected subcutaneously (S.C.). After four weeks the animals were injected S.C. with 200 micrograms of antigen mixed with an equal volume of IFA. Every four weeks animals were boosted with 100 micrograms of antigen. Seven days following each boost the animal was bled. Sera was generated by incubating the blood at 4°C for 12-24 hours followed by centrifugation.

25 Characterization of polyclonal antisera was carried out as follows. Ninety-six well plates were coated with antigen by incubating with 50 microliters (typically 1 microgram) at 4°C for 20 hrs. 250 microliters of BSA blocking buffer was added to the wells and incubated at room temperature for 2 hrs. Plates were washed 6 times with PBS/0.01% Tween. Rabbit sera was diluted in PBS and 50 microliters of  
30 diluted sera was added to each well and incubated at room temperature for 30 min.

Plates were washed as described above before addition of 50 microliters of goat anti-rabbit horse radish peroxidase (HRP) at a 1:10000 dilution and incubation at room temperature for 30 min. Plates were washed as described above and 100µl of TMB Microwell Peroxidase Substrate was added to each well. Following a 15 minute incubation in the dark at room temperature, the colorimetric reaction was stopped with 100µl 1N H<sub>2</sub>SO<sub>4</sub> and read immediately at 450 nm. Antisera showed strong reactivity to antigen L762P.

Immunohistochemical analysis using polyclonal antibodies against L762P demonstrated staining in all lung cancer samples tested, some light staining in the bronchiole epithelium of normal lung, tubule staining in kidney, light epithelial staining in colon and no staining in heart or liver.

In order to evaluate L773P protein expression in various tissues, immunohistochemistry (IHC) analysis was performed using an affinity purified L773P polyclonal antibody. Briefly, tissue samples were fixed in formalin solution for 12-24 hrs and embedded in paraffin before being sliced into 8 micron sections. Steam heat induced epitope retrieval (SHIER) in 0.1 M sodium citrate buffer (pH 6.0) was used for optimal staining conditions. Sections were incubated with 10% serum/PBS for 5 minutes. Primary antibody was added to each section for 25 minutes at indicated concentrations followed by 25 minute incubation with either anti-rabbit or anti-mouse biotinylated antibody. Endogenous peroxidase activity was blocked by three 1.5 minute incubations with hydrogen peroxidase. The avidin biotin complex/horse radish peroxidase (ABC/HRP) system was used along with DAB chromogen to visualize L773P expression. Slides were counterstained with hematoxylin to visualize cell nuclei. Using this approach, L773P protein was detected in 6/8 lung tumors, 4/6 normal lung samples (very light staining in some cases), 1/1 kidney samples (very light staining), 0/1 heart samples, 1/1 colon samples (very light staining) and 0/1 liver samples.

## EXAMPLE 7

PEPTIDE PRIMING OF MICE AND PROPAGATION OF CTL LINES

Immunogenic peptides from the lung cancer antigen L762P (SEQ ID  
5 NO: 161) for HLA-A2/K<sup>b</sup>-restricted CD8<sup>+</sup> T cells were identified as follows.

The location of HLA-A2 binding peptides within the lung cancer antigen  
L762P (SEQ ID NO: 161) was predicted using a computer program which predicts  
peptides sequences likely to be to HLA-A\*0201 by fitting to the known peptide  
binding motif for HLA-A\*0201 (Rupert *et al.* (1993) *Cell* 74:929; Rammensee *et al.*  
10 (1995) *Immunogenetics* 41:178-228). A series of 19 synthetic peptides corresponding  
to a selected subset of the predicted HLA-A\*0201 binding peptides was prepared as  
described above.

Mice expressing the transgene for human HLA A2/K<sup>b</sup> (provided by Dr L.  
Sherman, The Scripps Research Institute, La Jolla, CA) were immunized with the  
15 synthetic peptides, as described by Theobald *et al.*, *Proc. Natl. Acad. Sci. USA*  
92:11993-11997, 1995, with the following modifications. Mice were immunized with  
50µg of L762P peptide and 120µg of an I-A<sup>b</sup> binding peptide derived from hepatitis B  
virus protein emulsified in incomplete Freund's adjuvant. Three weeks later these mice  
were sacrificed and single cell suspensions prepared. Cells were then resuspended at 7  
20  $\times 10^6$  cells/ml in complete media (RPMI-1640; Gibco BRL, Gaithersburg, MD)  
containing 10% FCS, 2mM Glutamine (Gibco BRL), sodium pyruvate (Gibco BRL),  
non-essential amino acids (Gibco BRL),  $2 \times 10^{-5}$  M 2-mercaptoethanol, 50U/ml  
penicillin and streptomycin, and cultured in the presence of irradiated (3000 rads)  
L762P peptide- (5µg/ml) and 10mg/ml B<sub>2</sub>-microglobulin- (3 µg/ml) LPS blasts (A2  
25 transgenic spleens cells cultured in the presence of 7µg/ml dextran sulfate and 25µg/ml  
LPS for 3 days). After six days, cells ( $5 \times 10^5$ /ml) were restimulated with  $2.5 \times 10^6$ /ml  
peptide-pulsed irradiated (20,000 rads) EL4A2Kb cells (Sherman *et al.*, *Science*  
258:815-818, 1992) and  $5 \times 10^6$ /ml irradiated (3000 rads) A2/K<sup>b</sup>-transgenic spleen  
feeder cells. Cells were cultured in the presence of 10U/ml IL-2. Cells were  
30 restimulated on a weekly basis as described, in preparation for cloning the line.

Peptide-specific cell lines were cloned by limiting dilution analysis with irradiated (20,000 rads) L762P peptide-pulsed EL4 A2Kb tumor cells ( $1 \times 10^4$  cells/well) as stimulators and irradiated (3000 rads) A2/K<sup>b</sup>-transgenic spleen cells as feeders ( $5 \times 10^5$  cells/ well) grown in the presence of 10U/ml IL-2. On day 7, cells were  
 5 restimulated as before. On day 14, clones that were growing were isolated and maintained in culture.

Cell lines specific for the peptides L762P-87 (SEQ ID NO: 226; corresponding to amino acids 87-95 of SEQ ID NO: 161), L762P-145 (SEQ ID NO: 227; corresponding to amino acids 145-153 of SEQ ID NO: 161), L762P-585 (SEQ ID  
 10 NO: 228; corresponding to amino acids 585-593 of SEQ ID NO: 161), L762P-425 (SEQ ID NO: 229; corresponding to amino acids 425-433 of SEQ ID NO: 161), L762P(10)-424 (SEQ ID NO: 230; corresponding to amino acids 424-433 of SEQ ID NO: 161) and L762P(10)-458 (SEQ ID NO: 231; corresponding to amino acids 458-467 of SEQ ID NO: 161) demonstrated significantly higher reactivity (as measured by  
 15 percent specific lysis) against L762P peptide-pulsed EL4-A2K<sup>b</sup> tumor target cells than control peptide-pulsed EL4-A2/K<sup>b</sup> tumor target cells.

## EXAMPLE 8

### IDENTIFICATION OF CD4 IMMUNOGENIC T CELL EPITOPES DERIVED FROM 20 THE LUNG CANCER ANTIGEN L762P

CD4 T cell lines specific for the antigen L762P (SEQ ID NO: 161) were generated as follows.

A series of 28 overlapping peptides were synthesized that spanned approximately 50% of the L762P sequence. For priming, peptides were combined into  
 25 pools of 4-5 peptides, pulsed at 20 micrograms/ml into dendritic cells for 24 hours. The dendritic cells were then washed and mixed with positively selected CD4<sup>+</sup> T cells in 96 well U-bottomed plates. Forty cultures were generated for each peptide pool. Cultures were restimulated weekly with fresh dendritic cells loaded with peptide pools. Following a total of 3 stimulation cycles, cells were rested for an additional week and  
 30 tested for specificity to antigen presenting cells (APC) pulsed with peptide pools using

interferon-gamma ELISA and proliferation assays. For these assays, adherent monocytes loaded with either the relevant peptide pool or an irrelevant peptide were used as APC. T cell lines that appeared to specifically recognize L762P peptide pools both by cytokine release and proliferation were identified for each pool. Emphasis was placed on identifying T cells with proliferative responses. T cell lines that demonstrated either both L762P-specific cytokine secretion and proliferation, or strong proliferation alone were further expanded to be tested for recognition of individual peptides from the pools, as well as for recognition of recombinant L762P. The source of recombinant L762P was *E. coli*, and the material was partially purified and endotoxin positive.

These studies employed 10 micrograms of individual peptides, 10 or 2 micrograms of an irrelevant peptide, and 2 or 0.5 micrograms of either L762P protein or an irrelevant, equally impure, *E. coli* generated recombinant protein. Significant interferon-gamma production and CD4 T cell proliferation was induced by a number of L762P-derived peptides in each pool. The amino acid sequences for these peptides are provided in SEQ ID NO: 232-251. These peptides correspond to amino acids 661-680, 676-696, 526-545, 874-893, 811-830, 871-891, 856-875, 826-845, 795-815, 736-755, 706-725, 706-725, 691-710, 601-620, 571-590, 556-575, 616-635, 646-665, 631-650, 541-560 and 586-605, respectively, of SEQ ID NO: 161.

CD4 T cell lines that demonstrated specificity for individual L762P-derived peptides were further expanded by stimulation with the relevant peptide at 10 micrograms/ml. Two weeks post-stimulation, T cell lines were tested using both proliferation and IFN-gamma ELISA assays for recognition of the specific peptide. A number of previously identified T cells continued to demonstrate L762P-peptide specific activity. Each of these lines was further expanded on the relevant peptide and, following two weeks of expansion, tested for specific recognition of the L762P-peptide in titration experiments, as well as for recognition of recombinant *E. coli*-derived L762P protein. For these experiments, autologous adherent monocytes were pulsed with either the relevant L762P-derived peptide, an irrelevant mammaglobin-derived peptide, recombinant *E. coli*-derived L762P (approx. 50% pure), or an irrelevant *E. coli*-derived protein. The majority of T cell lines were found to show low affinity for the relevant

peptide, since specific proliferation and IFN-gamma ratios dramatically decreased as L762P peptide was diluted. However, four lines were identified that demonstrated significant activity even at 0.1 micrograms/ml peptide. Each of these lines (referred to as A/D5, D/F5, E/A7 and E/B6) also appeared to specifically proliferate in response to the *E. coli*-derived L762P protein preparation, but not in response to the irrelevant protein preparation. The amino acid sequences of the L762P-derived peptides recognized by these lines are provided in SEQ ID NO: 234, 249, 236 and 245, respectively. No protein specific IFN-gamma was detected for any of the lines. Lines A/D5, E/A7 and E/B6 were cloned on autologous adherent monocytes pulsed with the relevant peptide at 0.1 (A/D5 and E/A7) or 1 (D/F5) microgram/ml. Following growth, clones were tested for specificity for the relevant peptide. Numerous clones specific for the relevant peptide were identified for lines A/D5 and E/A7.

#### EXAMPLE 9

##### 15 PROTEIN EXPRESSION OF LUNG TUMOR-SPECIFIC ANTIGENS

###### a) Expression of L514S in *E. coli*

The lung tumor antigen L514S (SEQ ID NO: 89) was subcloned into the expression vector pE32b at NcoI and NotI sites, and transformed into *E. coli* using standard techniques. The protein was expressed from residues 3-153 of SEQ ID NO: 89. The expressed amino acid sequence and the corresponding DNA sequence are provided in SEQ ID NO: 252 and 253, respectively.

###### b) Expression of L762P

Amino acids 32-944 of the lung tumor antigen L762P (SEQ ID NO: 161), with a 6X His Tag, were subcloned into a modified pET28 expression vector, using kanamycin resistance, and transformed into BL21 CodonPlus using standard techniques. Low to moderate levels of expression were observed. The determined DNA sequence of the L762P expression construct is provided in SEQ ID NO: 254.

## EXAMPLE 10

IDENTIFICATION OF MHC CLASS II RESTRICTING ALLELE FOR L762PPEPTIDE-SPECIFIC RESPONSES

- 5           A panel of HLA mismatched antigen presenting cells (APC) were used to identify the MHC class II restricting allele for the L762P-peptide specific responses of CD4 T cell clones derived from lines that recognized L762P peptide and recombinant protein. Clones from two lines, AD-5 and EA-7, were tested as described below. The AD-5 derived clones were found to be restricted by the HLA-DRB-1101 allele, and an
- 10   EA-7 derived clone was found to be restricted by the HLA DRB-0701 or DQB1-0202 allele. Identification of the restriction allele allows targeting of vaccine therapies using the defined peptide to individuals that express the relevant class II allele. Knowing the relevant restricting allele will also enable clinical monitoring for responses to the defined peptide since only individuals that express the relevant allele will be monitored.
- 15           CD4 T cell clones derived from line AD-5 and EA-7 were stimulated on autologous APC pulsed with the specific peptide at 10 µg/ml, and tested for recognition of autologous APC (from donor D72) as well as against a panel of APC partially matched with D72 at class II alleles. Table 2 shows the HLA class typing of the APC tested. Adherent monocytes (generated by 2 hour adherence) from four different
- 20   donors, referred to as D45, D187, D208, and D326, were used as APC in these experiments. Autologous APC were not included in the experiment. Each of the APC were pulsed with the relevant peptide (5a for AD-5 and 3e for 3A-7) or the irrelevant mammoglobin peptide at 10 µg/ml, and cultures were established for 10,000 T cells and about 20,000 APC/well. As shown in Table 3, specific proliferation and cytokine
- 25   production could be detected only when partially matched donor cells were used as APC. Based on the MHC typing analysis, these results strongly suggest that the restricting allele for the L762-specific response of the AD-5 derived clones is HLA-DRB-1101 and for the EA-7 derived clone the restricting allele is HLA DRB-0701 or DQB1-0202.

Table 2 - HLA Typing of APC

DONOR	DR	DR	DQ	DQ
D72	B1-1101	B1-0701	B1-0202	B1-0301
D45	-3	-15	B1-0201	B1-0602
D187	-4	-15	-1	-7
D208	B1-1101	B1-0407	-3	-3
D326	B1-0301	B1-0701	B1-0202	B1-0201



Table 3 - I762P Peptide Responses Map to HLA DR Alleles

	AD-5																EA-7					
	A11		B10		C10		C11		E6		F1		F9		G8		G9		G10		G12	
	ProI	γ-IFN	ProI	γ-IFN	ProI	γ-IFN	ProI	γ-IFN	ProI	γ-IFN	ProI	γ-IFN	ProI	γ-IFN	ProI	γ-IFN	ProI	γ-IFN	ProI	γ-IFN	ProI	γ-IFN
Donor																						
D72 DR-0701, -1101, DQ-0202, -7	46		31		34		24		31		40	55		45		43		91		10		
D45 DR-3,-15, DQ-1, -0201	32	1.7	5.5	1.2	3.3	1	1.0	1.5	1.1	1.1	1.6	1.1	1.4	1.3	0.2	1.1	1.1	1.1	1.2	1.5	0.8	1.1
D187 DR-4, -15, DQ-1,-7	14	1.2	1.3	1	1.4	1.1	1.4	1.7	1.0	1.1	1.4	1.2	1.2	1.1	0.9	1	1.0	1	1.0	1.6	0.5	1
D208 DR-4, -1101, DQ-3	138	13	38	54	188	10	146	4.6	153	6.1	45.9	8.6	73.3	14.1	38.0	7.7	174.3	16.1	113.6	19.6	0.8	1
D326 DR-3, -0701, DQ-0202	0.7	4	0.3	1	0.3	1.4	1.0	2	0.8	1.1	0.3	1.1	0.7	1.1	0.6	1.2	0.4	1	1.2	5	14.1	6.8

## EXAMPLE 11

FUSION PROTEINS OF N-TERMINAL AND C-TERMINAL PORTIONS OF L763P

- In another embodiment, a *Mycobacterium tuberculosis*-derived polynucleotide, referred to as Ra12, is linked to at least an immunogenic portion of a polynucleotide of this invention. Ra12 compositions and methods for their use in enhancing expression of heterologous polynucleotide sequences are described in U.S. Patent Application 60/158,585, the disclosure of which is incorporated herein by reference in its entirety. Briefly, Ra12 refers to a polynucleotide region that is a subsequence of a *Mycobacterium tuberculosis* MTB32A nucleic acid. MTB32A is a serine protease of 32 KD molecular weight encoded by a gene in virulent and avirulent strains of *M. tuberculosis*. The nucleotide sequence and amino acid sequence of MTB32A have been described (for example, U.S. Patent Application 60/158,585; see also, Skeiky *et al.*, *Infection and Immun.* (1999) 67:3998-4007, incorporated herein by reference). Surprisingly, it was discovered that a 14 KD C-terminal fragment of the MTB32A coding sequence expresses at high levels on its own and remains as a soluble protein throughout the purification process. Moreover, this fragment may enhance the immunogenicity of heterologous antigenic polypeptides with which it is fused. This 14 KD C-terminal fragment of the MTB32A is referred to herein as Ra12 and represents a fragment comprising some or all of amino acid residues 192 to 323 of MTB32A.
- Recombinant nucleic acids which encode a fusion polypeptide comprising a Ra12 polypeptide and a heterologous lung tumor polypeptide of interest, can be readily constructed by conventional genetic engineering techniques. Recombinant nucleic acids are constructed so that, preferably, a Ra12 polynucleotide sequence is located 5' to a selected heterologous lung tumor polynucleotide sequence.
- It may also be appropriate to place a Ra12 polynucleotide sequence 3' to a selected heterologous polynucleotide sequence or to insert a heterologous polynucleotide sequence into a site within a Ra12 polynucleotide sequence.
- In addition, any suitable polynucleotide that encodes a Ra12 or a portion or other variant thereof can be used in constructing recombinant fusion polynucleotides comprising Ra12 and one or more lung tumor polynucleotides disclosed herein.

Preferred Ra12 polynucleotides generally comprise at least about 15 consecutive nucleotides, at least about 30 nucleotides, at least about 60 nucleotides, at least about 100 nucleotides, at least about 200 nucleotides, or at least about 300 nucleotides that encode a portion of a Ra12 polypeptide.

- 5 Ra12 polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a Ra12 polypeptide or a portion thereof) or may comprise a variant of such a sequence. Ra12 polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the biological activity of the encoded fusion polypeptide is not substantially diminished, relative to a fusion polypeptide comprising a native Ra12 polypeptide. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most preferably at least about 90% identity to a polynucleotide sequence that encodes a native Ra12 polypeptide or a portion thereof.

- Two specific embodiments of fusions between Ra12 and antigens of the present invention are described in this example.

A. N-Terminal Portion of L763P

- A fusion protein of full-length Ra12 and the N-terminal portion of L763P (referred to as L763P-N; amino acid residues 1-130 of SEQ ID NO: 159) was expressed as a single recombinant protein in *E. coli*. The cDNA for the N-terminal portion was obtained by PCR with a cDNA for the full length L763P and primers L763F3 (5' CGGCGAATTCATGGATTGGGGGACGCTGC; SEQ ID NO: 383) and 1763RV3 (5' CGGCCTCGAGTCAACCCTCTATCCGAACCTTCTGC; SEQ ID NO: 384). The PCR product with expected size was recovered from agarose gel, digested with restriction enzymes EcoRI and XhoI, and cloned into the corresponding sites in the expression vector pCRX1. The sequence for the fusion of full-length of Ra12 and L763P-N was confirmed by DNA sequencing. The determined cDNA sequence is provided in SEQ ID NO:351, with the corresponding amino acid sequence being provided in SEQ ID NO: 352).

B. C-Terminal Portion of L763P

A fusion protein of full-length Ra12 and the C-terminal portion of L763P (referred to as L763P-C; amino acid residues 100-262 of SEQ ID NO: 159) was expressed as a single recombinant protein in *E. coli*. The cDNA of the C-terminal portion of L763P was obtained by PCR with a cDNA for the full length of L763P and primers L763F4 (5' CGGCGAATTCCACGAACCACTCGCAAGTTCAG; SEQ ID NO: 385) and L763RV4 (5' CGGCTCGAG-TTAGCTTGGGCCTGTGATTGC; SEQ ID NO: 386). The PCR product with expected size was recovered from agarose gel, digested with restriction enzymes EcoRI and XhoI, and cloned into the corresponding sites in the expression vector pCRX1. The sequence for the fusion of full-length Ra12 and L763P-C was confirmed by DNA sequencing. The determined DNA sequence is provided in SEQ ID NO:353, with the corresponding amino acid sequence being provided in SEQ ID NO: 354.

The recombinant proteins described in this example are useful for the preparation of vaccines, for antibody therapeutics, and for diagnosis of lung tumors.

## EXAMPLE 12

### EXPRESSION IN *E. COLI* OF L762P HIS TAG FUSION PROTEIN

PCR was performed on the L762P coding region with the following primers:

Forward primer starting at amino acid 32.

PDM-278 5'ggagtacagcttcaagacaatggg 3' (SEQ ID NO:355) Tm 57°C.

Reverse primer including natural stop codon after amino acid 920, creating EcoRI site

PDM-280 5'ccatgggaattcattataataattttgtcc 3' (SEQ ID NO:356) Tm55°C.

The PCR product was digested with EcoRI restriction enzyme, gel purified and then cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and EcoRI restriction enzymes. The

correct construct was confirmed by DNA sequence analysis and then transformed into BL21 (DE3) pLys S and BL21 (DE3) CodonPlus RIL expression hosts.

The protein sequence of expressed recombinant L762P is shown in SEQ ID NO:357, and the DNA sequence is shown in SEQ ID NO:358.

5

### EXAMPLE 13

#### EXPRESSION IN *E. COLI* OF A L773PA HIS TAG FUSION PROTEIN

The L773PA coding region (encoding amino acids 2-71 of SEQ ID NO: 172) was PCR amplified using the following primers:

Forward primer for L773PA starting at amino acid 2:

PDM-299 5'tggcagccctcttcttcaagtggc 3' (SEQ ID NO:359) Tm63°C.

Reverse primer for L773PA creating artificial stop codon after amino acid 70:

PDM-355 5'cgccagaattcatcaaacaaatctgttagcacc 3' (SEQ ID NO:360) Tm62°C.

The resulting PCR product was digested with EcoRI restriction enzyme, gel purified and then cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco2I and EcoRI restriction enzymes. The correct construct was confirmed by DNA sequence analysis and transformed into BL21 (DE3) pLys S and BL21 (DE3) CodonPlus RIL expression hosts.

The protein sequence of expressed recombinant L773PA is shown in SEQ ID NO:361, and the DNA sequence is shown in SEQ ID NO:362.

25

### EXAMPLE 14

#### IDENTIFICATION OF EPITOPES DERIVED FROM LUNG TUMOR SPECIFIC POLYPEPTIDES

A series of peptides from the L773P amino acid sequence (SEQ ID NO: 172) were synthesized and used in *in vitro* priming experiments to generate peptide-specific CD4 T cells. These peptides were 20-mers that overlapped by 15 amino acids

- and corresponded to amino acids 1-69 of the L773P protein. This region has been demonstrated to be tumor-specific. Following three *in vitro* stimulations, CD4 T cell lines were identified that produced IFN $\gamma$  in response to the stimulating peptide but not the control peptide. Some of these T cell lines demonstrated recognition of recombinant
- 5 L773P and L773PA (tumor-specific region) proteins.

- To perform the experiments, a total of eleven 20-mer peptides (SEQ ID NOs: 363, 365 and 387-395) overlapping by 15 amino acids and derived from the N-terminal tumor-specific region of L773P (corresponding to amino acids 1-69 of SEQ ID NO:172) were generated by standard procedures. Dendritic cells were derived from
- 10 PBMC of a normal donor using GMCSF and IL-4 by standard protocol. Purified CD4 T cells were generated from the same donor as the dendritic cells using MACS beads and negative selection of PBMCs. Dendritic cells were pulsed overnight with the individual 20-mer peptides at a concentration of 10  $\mu$ g/ml. Pulsed dendritic cells were washed and plated at  $1 \times 10^4$ /well of a 96-well U-bottom plates, and purified CD4 cells were added
- 15 at  $1 \times 10^5$  well. Cultures were supplemented with 10 ng/ml IL-6 and 5 ng/ml IL-12, and incubated at 37°C. Cultures were re-stimulated as above on a weekly basis using as APC dendritic cells generated and pulsed as above, supplemented with 5 ng/ml IL-7 and 10  $\mu$ g/ml IL-2. Following 3 *in vitro* stimulation cycles, cell lines (each corresponding to one well) were tested for cytokine production in response to the stimulating peptide vs.
- 20 an irrelevant peptide.

- A small number of individual CD4 T cell lines (9/528) demonstrated cytokine release (IFN $\gamma$ ) in response to the stimulating peptide but not to control peptide. The CD4 T cell lines that demonstrated specific activity were restimulated on the appropriate L773P peptide and reassayed using autologous dendritic cells pulsed with
- 25 10  $\mu$ g/ml of the appropriate L773P peptide, an irrelevant control peptide, recombinant L773P protein (amino acids 2-364, made in *E. coli*), recombinant L773PA (amino acids 2-71, made in *E. coli*), or an appropriate control protein (L3E, made in *E. coli*). Three of the nine lines tested (1-3C, 1-6G, and 4-12B) recognized the appropriate L773P peptide as well as recombinant L773P and L773PA. Four of the lines tested (4-8A, 4-
- 30 8E, 4-12D, and 4-12E) recognized the appropriate L773P peptide only. Two of the lines tested (5-6F and 9-3B) demonstrated non-specific activity.

These results demonstrate that the peptide sequences MWQLFFKWLLSCCPGSSQI (amino acids 1-20 of SEQ ID NO: 172; SEQ ID NO:363) and GSSQIAAAASTQPEDDINTQ (amino acids 16-35 of SEQ ID NO: 172; SEQ ID NO: 365) may represent naturally processed epitopes of L773P, which are capable of stimulating human class II MHC-restricted CD4 T cell responses.

In subsequent studies, the above epitope mapping experiment was repeated using a different donor. Again, some of the resulting T cell lines were found to respond to peptide and recombinant protein. An additional peptide was found to be naturally processed. Specifically, purified CD4 cells were stimulated on a total of eleven 20-mer peptides overlapping by 15 amino acids (SEQ ID NO: 363, 387, 388, 365 and 389-395, respectively). The priming was carried out as described above, except that a peptide concentration of 0.5 ug/mL rather than 10 ug/mL was employed. In the initial screen of the cell lines 9 of the 528 lines released at least a three-fold greater level of IFN-gamma with stimulating peptide vs. control peptide. These 9 lines were restimulated on the appropriate peptide and then tested on dendritic cells pulsed with a titration of appropriate peptide (10 ug/mL, 1 ug/mL and 0.1 ug/mL), and 10 ug/mL of a control peptide. Six of the 9 lines recognized recombinant L773P as well as peptide. The six lines referred to as 1-1E, 1-2E, 1-4H, 1-6A, 1-6G and 2-12B recognized L773PA and the appropriate peptide. These results demonstrate that the peptides of SEQ ID NO: 363 and 387 represent naturally processed epitopes of L773P.

Using the procedures described above, CD4+ T cell responses were generated from PBMC of normal donors using dendritic cells pulsed with overlapping 20-mer peptides (SEQ ID NO: 396-419) spanning the L523S polypeptide sequence (SEQ ID NO: 176). A number of CD4+ T cells demonstrated reactivity with the priming peptides as well as with L523S recombinant protein, with the dominant reactivity of these lines being within the peptides 4, 7 and 21 (SEQ ID NO: 399, 402 and 416; corresponding to amino acids 30-39, 60-79 and 200-219, respectively, of SEQ ID NO: 176).

Epitopes within the scope of the invention include epitopes restricted by other class II MHC molecules. In addition, variants of the peptide can be produced wherein one or more amino acids are altered such that there is no effect on the ability of

the peptides to bind to MHC molecules, no effect on their ability to elicit T cell responses, and no effect on the ability of the elicited T cells to recognize recombinant protein.

5

## EXAMPLE 15

### SURFACE EXPRESSION OF L762P AND ANTIBODY EPITOPES THEREOF

Rabbits were immunized with full-length histidine-tagged L762P protein  
10 generated in *E. coli*. Sera was isolated from rabbits and screened for specific recognition of L762P in ELISA assays. One polyclonal serum, referred to as 2692L, was identified that specifically recognized recombinant L762P protein. The 2692L anti-L762P polyclonal antibodies were purified from the serum by affinity purification using L762P affinity columns. Although L762P is expressed in a subset of primary lung tumor  
15 samples, expression appears to be lost in established lung tumor cell lines. Therefore, to characterize surface expression of L762P, a retrovirus construct that expresses L762P was used to transduce primary human fibroblasts as well as 3 lung tumor cell lines (522-23, HTB, and 343T). Transduced lines were selected and expanded to examine L762P surface expression by FACS analysis. For this analysis, non-transduced and transduced  
20 cells were harvested using cell dissociation medium, and incubated with 10-50 micrograms/ml of either affinity purified anti-L762P or irrelevant antisera. Following a 30 minute incubation on ice, cells were washed and incubated with a secondary, FITC conjugated, anti rabbit IgG antibody as above. Cells were washed, resuspended in buffer with Propidium Iodide (PI) and examined by FACS using an Excalibur fluorescence  
25 activated cell sorter. For FACS analysis, PI-positive (i.e. dead/permeabilized cells) were excluded. The polyclonal anti-L762P sera specifically recognized and bound to the surface of L762P-transduced cells but not the non-transduced counterparts. These results demonstrate that L762P is localized to the cell surface of both fibroblasts as well as lung tumor cells.

30 To identify the peptide epitopes recognized by 2692L, an epitope mapping approach was pursued. A series of overlapping 19-21 mers (5 amino acid



- overlap) was synthesized that spanned the C terminal portion of L762P (amino acids 481-894 of SEQ ID NO: 161). In an initial experiment peptides were tested in pools. Specific reactivity with the L762P antiserum was observed with pools A, B, C, and E. To identify the specific peptides recognized by the antiserum, flat bottom 96 well
- 5 microtiter plates were coated with individual peptides at 10 microgram/ml for 2 hours at 37 °C. Wells were then aspirated and blocked with phosphate buffered saline containing 5% (w/v) milk for 2 hours at 37 °C, and subsequently washed in PBS containing 0.1% Tween 20 (PBST). Purified rabbit anti-L762P serum 2692L was added at 200 or 20 ng/well to triplicate wells in PBST and incubated overnight at room temperature. This
- 10 was followed by washing 6 times with PBST and subsequently incubating with HRP-conjugated donkey anti rabbit IgG (H+L)Affinipure F(ab') fragment at 1:2,000 for 60 minutes. Plates were then washed, and incubated in tetramethyl benzidine substrate. Reactions were stopped by the addition of 1N sulfuric acid and plates were read at 450/570 nm using an ELISA plate reader.
- 15                   The resulting data, presented in Table 4 below, demonstrates that the L762P antisera recognized at least 6 distinct peptide epitopes from the 3' half of L762P.

Table 4

Peptide (starting amino acid of L762P)	pool	ELISA activity (OD 450-570)	
		200 ng polyclonal serum	20 ng polyclonal serum
A (481)	A	1.76	1.0
B (495)	A	0.14	.06
C (511)	E	0.47	0.18
D (526)	E	0.11	0.09
E (541)	A	0.11	0.04
F (556)	A	0.04	0.02
G (571)	A	0.06	0.02
H (586)	B	0.1	0.03
I (601)	B	0.25	0.06
J (616)	B	0.1	0.03
K (631)	E	0.1	0.08
L (646)	B	0.28	0.12
M (661)	B	0.14	0.03
N (676)	C	0.12	0.1
O (691)	C	1.1	0.23
P (706)	C	0.1	0.03
Q (721)	C	0.11	0.05
R (736)	E	0.12	0.04
S (751)	C	0.15	0.06
U (781)	D	0.12	0.06
V (795)	F	0.07	0.05
X (826)	D	0.1	0.03
Y (841)	D	0.17	0.07
Z (856)	D	0.16	0.08
AA (871)	F	0.17	0.05
BB (874)	F	0.14	0.11
No peptide		0.15	0.045

- Individual peptides were identified from each of the pools, and
- 5 additionally a weak reactivity was identified with peptide BB from pool F. The relevant peptide epitopes are summarized in the Table 5 below. The amino acid sequences for peptides BB, O, L, I, A and C are provided in SEQ ID NO: 376-381, respectively, with the corresponding cDNA sequences being provided in SEQ ID NO: 373, 370, 372, 374, 371 and 375, respectively.

Table 5

Peptide	Nucleotides of L762P	Amino acids of L762P	Sequence	pool	<u>ELISA activity</u> <u>(OD 450-570)</u>	
					200 ng	20 ng
A	1441-1500	481-500	SRSSGTGDIFQQHIQLEST	A	1.76	1.0
C	1531-1590	511-530	KNTVTVDNTVGNDTMFLVTW	E	0.47	0.18
I	1801-1860	601-620	AVPPATVEAFVERDSLHFPH	B	0.25	0.06
L	1936-1955	646-665	PETGDPVTILRLDDGAGADV	B	0.28	0.12
O	2071-2130	691-710	VNHSPSISTPAHSIPGSHAMIL	C	1.1	0.23
BB	2620-2679	874-893	LQSAVSNIAQAPLFPPNSD	F	0.14	0.11
None	-	-	-	-	0.15	0.05

## EXAMPLE 16

5        DETECTION OF ANTIBODIES AGAINST LUNG TUMOR ANTIGENS  
IN PATIENT SERA

Antibodies specific for the lung tumor antigens L773PA (SEQ ID NO:361), L514S (SEQ ID NO:155 and 156), L523S (SEQ ID NO:176), L762P (SEQ ID NO:161) and L763P (SEQ ID NO:159) were shown to be present in effusion fluid or sera of lung cancer patients but not in normal donors. More specifically, the presence of antibodies against L773PA, L514S, L523S, L762P and L763P in effusion fluid obtained from lung cancer patients and in sera from normal donors was detected by ELISA using recombinant proteins and HRP-conjugated anti-human Ig. Briefly, each protein (100 ng) was coated in 96-well plate at pH 9.5. In parallel, BSA (bovine serum albumin) was also coated as a control protein. The signals ([S], absorbance measured at 405 nm) against BSA ([N]) were determined. The results of these studies are shown in Table 6, wherein - represents [S]/[N] < 2; +/- represents [S]/[N] > 2; ++ represents [S]/[N] > 3; and +++ represents [S]/[N] > 5.

20

Table 6 – Detection of Antibodies against Lung Tumor Antigens

	L514S	L523S	L762P	L763P	L773PA
Effusion fluid					
#1	+++	++	++	-	++
#2	-	-	+/-	++	+/-
#3	-	-	-	-	+/-
#4	+/-	++	+/-	-	+/-
#5	+/-	+++	+/-	+/-	++
#7	-	+/-	-	-	+/-
#8	-	+++	-	-	++
#10	-	++	+/-	+/-	-
#11	+/-	++	++	-	++
#12	+++	+/-	-	+/-	+/-
#13	-	+/-	-	-	+/-
#14	-	+++	+/-	+/-	++
#15	+/-	++	+/-	-	++
#17	-	+/-	-	-	+/-
#18	-	++	-	-	-
#19	-	+/-	-	-	+/-
#20	+/-	+/-	+/-	-	+/-
Normal sera					
#21	-	+/-	-	-	-
#22	-	-	-	-	-
#23	-	-	-	-	+/-
#24	-	+/-	-	-	-
#25	+/-	+/-	-	-	+/-

- Using Western blot analyses, antibodies against L523S were found to be present in 3 out of 4 samples of effusion fluid from lung cancer patients, with no L523S antibodies being detected in the three samples of normal sera tested.

#### EXAMPLE 17

#### 10 EXPRESSION IN *E. COLI* OF A L514S HIS TAG FUSION PROTEIN

PCR was performed on the L514S-13160 coding region with the following primers:

Forward primer PDM-278 5' cacactagtgtccgcgtgcgccctac 3' (SEQ ID NO:421) Tm 67°C.

Reverse primer PDM-280 5' catgagaattcatcacatgcccttgaaggctccc 3' (SEQ ID NO:422) Tm 66°C.

5 The PCR conditions were as follows:

10µl 10X Pfu buffer

1.0µl 10mM dNTPs

2.0µl 10µM each primer

83µl sterile water

10 1.5µl Pfu DNA polymerase (Stratagene, La Jolla, CA)

50ng DNA

96°C for 2 minutes, 96°C for 20 seconds, 66°C for 15 seconds, 72°C for 1 minute with 40 cycles and then 72°C for 4 minutes.

15 The PCR product was digested with EcoRI restriction enzyme, gel purified and then cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and EcoRI restriction enzymes. The correct construct was confirmed by DNA sequence analysis and then transformed into BL21 CodonPlus (Stratagene, La Jolla, CA) cells for expression.

20 The amino acid sequence of expressed recombinant L514S is shown in SEQ ID NO:423, and the DNA coding region sequence is shown in SEQ ID NO:424.

## EXAMPLE 18

### EXPRESSION IN *E. COLI* OF A L523S HIS TAG FUSION PROTEIN

25 PCR was performed on the L523S coding region with the following primers:

Forward primer PDM-414 5' aacaaactgtatatcggaacctcagcgagaa 3' (SEQ ID NO:425) Tm 62°C.

30 Reverse primer PDM-415 5' ccatagaattcattactccgtcttgactgagg 3' (SEQ ID NO:426) Tm 62°C.

The PCR conditions were as follows:

10µl 10X Pfu buffer

1.0µl 10mM dNTPs

2.0µl 10µM each primer

5 83µl sterile water

1.5µl Pfu DNA polymerase (Stratagene, La Jolla, CA)

50ng DNA

96°C for 2 minutes, 96°C for 20 seconds, 62°C for 15 seconds, 72°C for 4 minutes with 40 cycles and then 72°C for 4 minutes.

10 The PCR product was digested with EcoRI restriction enzyme, gel purified and then cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and EcoRI restriction enzymes. The correct construct was confirmed by DNA sequence analysis and then transformed into BL21 CodonPlus (Stratagene, La Jolla, CA) cells for expression.

15 The amino acid sequence of expressed recombinant L523S is shown in SEQ ID NO:427, and the DNA coding region sequence is shown in SEQ ID NO:428.

#### EXAMPLE 19

##### 20 EXPRESSION IN *E. COLI* OF A L762PA HIS TAG FUSION PROTEIN

PCR was performed on the L762PA coding region (L762PA is missing the signal sequence, the C-terminal transmembrane domain and the cytoplasmic tail) with the following primers:

25 Forward primer PDM-278 5'ggagtacagcttcaagacaatggg 3' (SEQ ID NO:355) Tm 57°C.

Reverse primer PDM-279 5'ccatggaattcattattcaatataagataatctc 3' (SEQ ID NO:429) Tm 56°C.

The PCR conditions were as follows:

10µl 10X Pfu buffer

30 1.0µl 10mM dNTPs

2.0μl 10μM each primer

83μl sterile water

1.5μl Pfu DNA polymerase (Stratagene, La Jolla, CA)

50ng DNA

5                    96°C for 2 minutes, 96°C for 20 seconds, 55°C for 15 seconds, 72°C for  
5 minutes with 40 cycles and then 72°C for 4 minutes.

The PCR product was digested with EcoRI restriction enzyme, gel  
purified and then cloned into pPDM His, a modified pET28 vector with a His tag in  
10    frame, which had been digested with Eco2I and EcoRI restriction enzymes. The  
correct construct was confirmed by DNA sequence analysis and then transformed into  
BL21 pLys S (Novagen, Madison, WI) cells for expression.

The amino acid sequence of expressed recombinant L762PA is shown in  
SEQ ID NO:430, and the DNA coding region sequence is shown in SEQ ID NO:431.

15

#### EXAMPLE 20

##### EXPRESSION IN *E. COLI* OF A L773P HIS TAG FUSION PROTEIN

PCR was performed on the L773P coding region with the following  
primers:

20                    Forward primer PDM-299    5' tggcagccctcttcttcaagtggc 3' (SEQ ID  
NO:359) Tm 63°C.

Reverse primer PDM-300    5' cgcctgctcgagtcattaattatcatcagaaaatgg 3'  
(SEQ ID NO:432) TM 63°C.

The PCR conditions were as follows:

25                    10μl 10X Pfu buffer

1.0μl 10mM dNTPs

2.0μl 10μM each primer

83μl sterile water

1.5μl Pfu DNA polymerase (Stratagene, La Jolla, CA)

30                    50ng DNA

96°C for 2 minutes, 96°C for 20 seconds, 63°C for 15 seconds, 72°C for 2 minutes 15 seconds with 40 cycles and then 72°C for 4 minutes.

The PCR product was digested with EcoRI restriction enzyme, gel purified and then cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and EcoRI restriction enzymes. The correct construct was confirmed by DNA sequence analysis and then transformed into BL21 pLys S (Novagen, Madison, WI) and BL21 CodonPlus (Stratagene, La Jolla, CA) cells for expression.

The amino acid sequence of expressed recombinant L773P is shown in SEQ ID NO:433, and the DNA coding region sequence is shown in SEQ ID NO:434.

## EXAMPLE 21

### CLONING AND SEQUENCING OF A T-CELL RECEPTOR CLONE FOR THE LUNG SPECIFIC ANTIGEN L762P

T cell receptor (TCR) alpha and beta chains from a CD4 T cell clone specific for the lung specific antigen L762P were cloned and sequence. Basically, total mRNA from  $2 \times 10^6$  cells from CTL clone 4H6 was isolated using Trizol reagent and cDNA was synthesized using Ready-to go kits (Pharmacia). To determine Valpha and Vbeta sequences of this clone, a panel of Valpha and Vbeta subtype specific primers was synthesized and used in RT-PCR reactions with cDNA generated from each of the clones. The RT-PCR reactions demonstrated that each of the clones expressed a common Vbeta sequence that corresponded to the Vbeta8 subfamily and a Valpha sequence that corresponded to the Valpha8 subfamily. To clone the full TCR alpha and beta chains from clone 4H6, primers were designed that spanned the initiator and terminator-coding TCR nucleotides. The primers were as follows:

forward primer for TCR Valpha8 5'  
ggatccgccgcccacatgacatccattcgagctgta 3' (SEQ ID NO:435; has a BamHI site inserted);



Kozak reverse primer for TCR Valpha8 (antisense) 5'  
gtcgactcagctggaccacagccgcag 3' (SEQ ID NO:436; has a SalI site inserted plus  
the TCR alpha constant sequence);

forward primer for TCR Vbeta8 (sense) 5'  
5 ggatccgccgcaccatggactcctggacctctgct 3' (SEQ ID NO:437; has a BamHI site  
inserted); and

Kozak reverse primer for TCR Vbeta 5' gtcgactcagaaatcctttctcttgac 3'  
(SEQ ID NO:438; has a SalI site inserted plus the TCR beta constant sequence).

Standard 35 cycle RT-PCR reactions were established using the cDNA synthesized  
10 from the CTL clone and the above primers utilizing the proofreading thermostable  
polymerase, PWO (Roche). The resultant PCR band, about 850 bp for Valpha and  
about 950 for Vbeta, was ligated into a PCR blunt vector (Invitrogen) and transformed  
into *E. coli*. *E. coli* transformed with plasmids having full-length alpha and beta chains  
were identified.. Large scale preparations of the corresponding plasmids were  
15 generated, and these plasmids were sequenced. The Valpha sequence (SEQ ID  
NO:439) was shown by nucleotide sequence alignment to be homologous to Valpha8.1,  
while the Vbeta sequence (SEQ ID NO:440) was shown by nucleotide sequence  
alignment to be homologous to Vbeta8.2.

20

## EXAMPLE 22

### RECOMBINANT EXPRESSION OF FULL LENGTH L762P IN MAMMALIAN CELLS

Full length L762P cDNA was subcloned into the mammalian expression  
25 vectors VR1012 and pCEP4 (Invitrogen). Both expression vectors had previously been  
modified to contain a FLAG epitope tag. These constructs were transfected into  
HEK293 and CHL-1 cells (ATCC) using Lipofectamine 2000 reagent (Gibco). Briefly,  
both the HEK and CHL-1 cells were plated at a density of 100,000 cells/ml in DMEM  
(Gibco) containing 10% FBS (Hyclone) and grown overnight. The following day, 4μl  
30 of Lipofectamine 2000 was added to 100μl of DMEM containing no FBS and incubated

for 5 minutes at room temperature. The Lipofectamine/DMEM mixture was then added to 1 µg of L762P Flag/pCEP4 or L762P Flag/VR1012 plasmid DNA resuspended in 100 µl DMEM and incubated for 15 minutes at room temperature. The Lipofectamine/DNA mix was then added to the HEK293 and CHL-1 cells and  
5 incubated for 48-72 hours at 37°C with 7% CO<sub>2</sub>. Cells were rinsed with PBS, then collected and pelleted by centrifugation. L762P expression was detected in the transfected HEK293 and CHL-1 cell lysates by Western blot analysis and was detected on the surface of transfected HEK cells by flow cytometry analysis.

For Western blot analysis, whole cell lysates were generated by  
10 incubating the cells in Triton-X100 containing lysis buffer for 30 minutes on ice. Lysates were then cleared by centrifugation at 10,000 rpm for 5 minutes at 4°C. Samples were diluted with SDS-PAGE loading buffer containing beta-mercaptoethanol, then boiled for 10 minutes prior to loading the SDS-PAGE gel. The protein was transferred to nitrocellulose and probed using 1 µg/ml purified anti-L762P rabbit  
15 polyclonal sera (lot #690/73) or non-diluted anti-L762P mAb 153.20.1 supernatant. Blots were revealed using either goat anti-rabbit Ig coupled to HRP or goat anti-mouse Ig coupled to HRP followed by incubation in ECL substrate.

For flow cytometric analysis, cells were washed further with ice cold staining buffer (PBS+1%BSA +Azide). Next, the cells were incubated for 30 minutes  
20 on ice with 10 µg/ml of purified anti-L762P polyclonal sera (lot #690/73) or a 1:2 dilution of anti-L762P mAb 153.20.1 supernatant. The cells were washed 3 times with staining buffer and then incubated with a 1:100 dilution of goat anti-rabbit Ig(H+L)-FITC or goat anti-mouse Ig(H+L)-FITC reagent (Southern Biotechnology) for 30 minutes on ice. After 3 washes, the cells were resuspended in staining buffer containing  
25 propidium iodide (PI), a vital stain that allows for the exclusion of permeable cells, and analyzed by flow cytometry.

## EXAMPLE 23

GENERATION OF POLYCLONAL ANTIBODIES TO LUNG TUMOR ANTIGENS

Three lung antigens, L523S (SEQ ID NO:176), L763P (SEQ ID NO:159) and L763 peptide #2684 (SEQ ID NO:441), were expressed and purified for use in  
5 antibody generation.

L523S and L763P were expressed in an *E. coli* recombinant expression system and grown overnight in LB Broth with the appropriate antibiotics at 37°C in a shaking incubator. The next morning, 10 ml of the overnight culture was added to 500 ml of 2x YT with the appropriate antibiotics in a 2L-baffled Erlenmeyer flask. When  
10 the optical density of the culture reached 0.4-0.6 at 560 nanometers, the cells were induced with IPTG (1 mM). Four hours after induction with IPTG, the cells were harvested by centrifugation.

The cells were then washed with phosphate buffered saline and centrifuged again. The supernatant was discarded and the cells were either frozen for  
15 future use or immediately processed. Twenty milliliters of lysis buffer was added to the cell pellets and vortexed. To break open the *E. coli* cells, this mixture was then run through a french press at a pressure of 16,000 psi. The cells were then centrifuged again and the supernatant and pellet were checked by SDS-PAGE for the partitioning of the recombinant protein.

For proteins that localized to the cell pellet, the pellet was resuspended in  
20 10 mM Tris pH 8.0, 1% CHAPS and the inclusion body pellet was washed and centrifuged again. This procedure was repeated twice more. The washed inclusion body pellet was solubilized with either 8M urea or 6M guanidine HCl containing 10 mM Tris pH 8.0 plus 10 mM imidazole. The solubilized protein was added to 5 ml of nickel-  
25 chelate resin (Qiagen) and incubated for 45 minutes to 1 hour at room temperature with continuous agitation.

After incubation, the resin and protein mixture was poured through a disposable column and the flow through was collected. The column was then washed  
30 with 10-20 column volumes of the solubilization buffer. The antigen was then eluted from the column using 8M urea, 10 mM Tris pH 8.0 and 300 mM imidazole and

collected in 3 ml fractions. A SDS-PAGE gel was run to determine which fractions to pool for further purification.

As a final purification step, a strong anion exchange resin, in this case Hi-Prep Q (Biorad), was equilibrated with the appropriate buffer and the pooled fractions from above were loaded onto the column. Each antigen was eluted off the column with an increasing salt gradient. Fractions were collected as the column was run and another SDS-PAGE gel was run to determine which fractions from the column to pool.

The pooled fractions were dialyzed against 10 mM Tris pH 8.0. The release criteria were purity as determined by SDS-PAGE or HPLC, concentration as determined by Lowry assay or Amino Acid Analysis, identity as determined by amino terminal protein sequence, and endotoxin level was determined by the Limulus (LAL) assay. The proteins were then put in vials after filtration through a 0.22-micron filter and the antigens were frozen until needed for immunization.

The L763 peptide #2684 was synthesized and conjugated to KLH and froze until needed for immunization.

The polyclonal antisera were generated using 400 micrograms of each lung antigen combined with 100 micrograms of muramyl dipeptide (MDP). An equal volume of Incomplete Freund's Adjuvant (IFA) was added and then mixed and injected subcutaneously (S.C.) into a rabbit. After four weeks, the rabbit was S.C. boosted with 200 micrograms of antigen mixed with an equal volume of IFA. Thereafter the rabbit was I.V. boosted with 100 micrograms of antigen. The animal was bled seven days following each boost. The blood was then incubated at 4°C for 12-24 hours followed by centrifugation to generate the sera.

The polyclonal antisera were characterized using 96 well plates coated with antigen and incubated with 50 microliters (typically 1 microgram/microliter) of the polyclonal antisera at 4°C for 20 hours. Basically, 250 microliters of BSA blocking buffer was added to the wells and incubated at room temperature for 2 hours. Plates were washed 6 times with PBS/0.1% Tween. The rabbit sera were diluted in PBS/0.1% Tween/0.1%BSA. 50 microliters of diluted sera was added to each well and incubated

at room temperature for 30 minutes. The plates were washed as described above, and then 50 microliters of goat anti-rabbit horseradish peroxidase (HRP) at a 1:10000 dilution was added and incubated at room temperature for 30 minutes.

- The plates were washed as described above, and 100 microliters of TMB
- 5 Microwell Peroxidase Substrate was added to each well. Following a 15-minute incubation in the dark at room temperature, the colorimetric reaction was stopped with 100 microliters of 1N H<sub>2</sub>SO<sub>4</sub> and read immediately at 450 nm. All the polyclonal antibodies showed immunoreactivity to the appropriate antigen. Tables 7-9 show the antibody reactivity of rabbit antisera in serial dilution to the three lung antigens, L523S,
- 10 L763P and L763 peptide #2684. The first column shows the antibody dilutions. The columns "Pre-immune sera" indicate ELISA data for two experiments using pre-immune sera. These results are averaged in the fourth column. The columns "anti-L523S, L763P or #2684" indicate ELISA data for two experiments using sera from rabbits immunized as described in this Example, using the respective antigen, referred
- 15 to as either L523S, L763P or #2684 in the tables.

Table 7

Antibody dilution	Pre-immune sera (1)	Pre-immune sera (2)	Average	Anti-L523S (1)	Anti-L523S (2)	Average
1:1000	0.14	0.14	0.14	2.36	2.37	2.37
1:2000	0.12	0.10	0.11	2.29	2.23	2.26
1:4000	0.10	0.09	0.10	2.11	2.17	2.14
1:8000	0.09	0.09	0.09	1.98	2.00	1.99
1:16000	0.09	0.09	0.09	1.73	1.76	1.75
1:32000	0.09	0.09	0.09	1.35	1.40	1.37
1:64000	0.09	0.11	0.10	0.94	0.98	0.96
1:128000	0.09	0.08	0.08	0.61	0.61	0.61
1:256000	0.08	0.08	0.08	0.38	0.38	0.38
1:512000	0.09	0.08	0.08	0.24	0.25	0.25
1:1024000	0.08	0.08	0.08	0.17	0.17	0.17
1:2048000	0.08	0.08	0.08	0.14	0.13	0.13

Table 8

Antibody dilution	Pre-immune sera (1)	Pre-immune sera (2)	Average	Anti-L763P (1)	Anti-L763P (2)	Average
1:1000	0.09	0.11	0.10	1.97	1.90	1.93
1:2000	0.07	0.07	0.07	1.86	1.84	1.85
1:4000	0.06	0.06	0.06	1.82	1.81	1.81
1:8000	0.06	0.06	0.06	1.83	1.81	1.82
1:16000	0.06	0.05	0.06	1.79	1.74	1.76
1:32000	0.06	0.06	0.06	1.56	1.51	1.53
1:64000	0.06	0.05	0.05	1.35	1.34	1.35
1:128000	0.05	0.05	0.05	1.01	0.98	0.99
1:256000	0.06	0.05	0.05	0.69	0.70	0.70
1:512000	0.06	0.05	0.05	0.47	0.44	0.46
1:1024000	0.06	0.05	0.06	0.27	0.27	0.27
1:2048000	0.05	0.05	0.05	0.16	0.15	0.16

Table 9

Antibody dilution	Pre-immune sera (1)	Pre-immune sera (2)	Average	Anti- #2684 (1)	Anti- #2684 (2)	Average
1:1000	0.07	0.07	0.07	2.10	2.00	2.05
1:2000	0.07	0.06	0.06	1.95	1.96	1.95
1:4000	0.06	0.06	0.06	1.77	1.82	1.79
1:8000	0.06	0.06	0.06	1.79	1.81	1.80
1:16000	0.06	0.06	0.06	1.54	1.50	1.52
1:32000	0.06	0.06	0.06	1.27	1.20	1.24
1:64000	0.06	0.06	0.06	0.85	0.82	0.83
0	0.06	0.06	0.06	0.06	0.06	0.06

5                      Tables 10-12 show the affinity purification of the respective antibodies to the three lung antigens, L523S, L763P and L763 peptide #2684.

Table 10

Antibody conc. (µg/ml)	Affinity pure (salt peak)	Affinity pure (salt peak)	Average	Affinity pure (acid peak)	Affinity pure (acid peak)	Average
1.0	2.38	2.35	2.36	2.25	2.31	2.28
0.5	2.24	2.22	2.23	2.19	2.18	2.18
0.25	2.05	2.09	2.07	2.01	2.03	2.02
0.13	1.70	1.81	1.75	1.74	1.74	1.74
0.063	1.44	1.44	1.44	1.43	1.38	1.40
0.031	1.05	1.05	1.05	0.99	0.99	0.99
0.016	0.68	0.67	0.68	0.65	0.64	0.64
0.0078	0.43	0.42	0.42	0.39	0.39	0.39
0.0039	0.27	0.26	0.27	0.24	0.26	0.25
0.0020	0.18	0.20	0.19	0.19	0.18	0.19
0.0010	0.13	0.14	0.13	0.13	0.14	0.13
0.00	0.11	0.12	0.11	0.10	0.12	0.11

Table 11

Antibody dilution	Affinity pure	Affinity pure	Average
1:1000	1.64	1.77	1.70
1:2000	1.59	1.76	1.68
1:4000	1.48	1.62	1.55
1:8000	1.35	1.43	1.39
1:16000	1.09	1.19	1.14
1:32000	0.81	0.89	0.85
1:64000	0.55	0.58	0.56
1:128000	0.31	0.35	0.33
1:256000	0.18	0.20	0.19
1:512000	0.11	0.12	0.11
1:1024000	0.07	0.07	0.07
1:2048000	0.06	0.06	0.06

Table 12

5

Antibody conc. (µg/ml)	Affinity pure	Affinity pure	Average
1.0	2.00	2.02	2.01
0.5	2.01	1.93	1.97
0.25	1.84	1.83	1.84
0.13	1.80	1.83	1.81
0.06	1.39	1.60	1.50
0.03	1.33	1.35	1.34
0.02	0.94	0.93	0.94
0.00	0.06	0.06	0.06

## EXAMPLE 24

FULL-LENGTH cDNA SEQUENCE ENCODING L529S

10

The isolation of a partial sequence (SEQ ID NO:106) for lung antigen L529S was previously provided in Example 2. This partial sequence was used as a



query to identify potential full length cDNA and protein sequences by searching against publicly available databases. The predicted full-length cDNA sequence for the isolated cloned sequence of SEQ ID NO:106 is provided in SEQ ID NO:442. The deduced amino acid sequence of the antigen encoded by SEQ ID NO:442 is provided in SEQ ID NO:443. It was previously disclosed in Example 2 that L529S shows similarity to connexin 26, a gap junction protein.

### EXAMPLE 25

#### EXPRESSION IN MEGATERIUM OF A HISTIDINE TAG-FREE L523S FUSION PROTEIN

PCR was performed on the L523S coding region with the following primers:

Forward primer PDM-734 5' caatcagcatgcacacaaactgtatatcggaac 3' (SEQ ID NO:444) Tm 63°C.

Reverse primer PDM-735 5' cgteaagatcttcattactccgtcttgac 3' (SEQ ID NO:445) Tm 60°C.

The PCR conditions were as follows:

10µl 10X Pfu buffer  
1.0µl 10mM dNTPs  
2.0µl 10µM each primer  
83µl sterile water  
1.5µl Pfu DNA polymerase (Stratagene, La Jolla, CA)  
50ng DNA  
96°C for 2 minutes, 96°C for 20 seconds, 62°C for 15 seconds, 72°C for 4 minute with 40 cycles and then 72°C for 4 minutes.

The PCR product was digested with SphI and BglII restriction enzymes, gel purified and then cloned into pMEG-3, which had been digested with SphI and BglII restriction enzymes. The correct construct was confirmed by DNA sequence analysis and then transformed into Megaterium cells for expression.

The amino acid sequence of expressed recombinant L523S is shown in SEQ ID NO:446, and the DNA coding region sequence is shown in SEQ ID NO:447.

5

#### EXAMPLE 26

#### EXPRESSION IN *E. COLI* OF A HISTIDINE TAG-FREE L523S FUSION PROTEIN

PCR was performed on the L552S coding region with the following primers:

10 Forward primer PDM-733 5' cgtactagcatatgaacaaactgtatatcggaac 3'  
(SEQ ID NO:448) Tm 64°C.

Reverse primer PDM-415 5' ccatagaattcattactccgtcttgactgagg 3' (SEQ  
ID NO:426) Tm 62°C.

The PCR conditions were as follows:

15 10µl 10X Pfu buffer  
1.0µl 10mM dNTPs  
2.0µl 10µM each primer  
83µl sterile water  
1.5µl Pfu DNA polymerase (Stratagene, La Jolla, CA)  
20 50ng DNA

96°C for 2 minutes, 96°C for 20 seconds, 62°C for 15 seconds, 72°C for  
4 minute with 40 cycles and then 72°C for 4 minutes.

25 The PCR product was digested with NdeI and EcoRI restriction  
enzymes, gel purified and then cloned into pPDM, a modified pET28 vector, which had  
been digested with NdeI and EcoRI restriction enzymes. The correct construct was  
confirmed by DNA sequence analysis and then transformed into BLR pLys S and HMS  
174 pLys S cells for expression.

The amino acid sequence of expressed recombinant L523S is shown in  
SEQ ID NO:449, and the DNA coding region sequence is shown in SEQ ID NO:450.

30

EXAMPLE 27

## EPIOTOPE-ANALYSIS OF L514S AND L523S-SPECIFIC ANTIBODIES

5           Peptides of candidate antigens can be used for the evaluation of antibody responses in both preclinical and clinical studies. These data allow one to further confirm the antibody response against a certain candidate antigen. Protein-based ELISA with and without competitive peptides and peptide-based ELISA can be used to evaluate these antibody responses. Peptide ELISA is especially useful since it can  
10 further exclude the false positive of the antibody titer observed in protein-based ELISA as well as to provide the simplest assay system to test antibody responses to candidate antigens. In this example, data was obtained using both L514S- and L523S-peptides that show that individual cancer patients produce L514S- and L523S-specific antibodies. The L514S-specific antibodies recognize primarily the following epitope of  
15 L514S:

aa86-110:       LGKEVRDAKITPEAFEKLGFPAAKE       (SEQ ID NO:451).

This epitope is the common epitope in humans. A rabbit antibody specific for L514S recognizes two addition epitopes of L514S:

20           (1) aa21-45:       KASDGDYYTLAVPMGDPMDGISVA (SEQ ID NO:452)

(2) aa121-135:       PDRDVLNTHQLNPKVK       (SEQ ID NO:453)

It was further found that the SEQ ID NO:452 is common to both L514S isoforms, L514S-13160 and L514S-13166, whereas the other epitopes, SEQ ID NO:451  
25 and SEQ ID NO:453, are probably specific to the isoform, L514S-13160.

The L523S-specific antibodies recognize primarily the following epitope of L523S:

aa440-460:       KIAPAEAPDAKVRMVIITGP       (SEQ ID NO:454).

This epitope is the common epitope in humans. A rabbit antibody  
30 specific for L523S recognizes two other epitopes:

- (1) aa156-175 PDGAAQQNNNPLQQPRG (SEQ ID  
NO:455)
- (2) aa326-345: RTITVKGNVETCAKAEEEEIM (SED ID  
NO:456)

5

In further studies, it was determined by peptide based ELISAs that eight additional epitopes of L523S were recognized by L523S-specific antibodies:

- (1) aa40-59 AFVDCPDESWALKAEALS (SEQ ID  
NO:457)
- (2) aa80-99: IRKLQIRNIPPHLQWEVLDS (SED ID  
NO:458)
- (3) aa160-179: AQQNPLQQPRGRRGLGQRGS (SEQ ID  
NO:459)
- (4) aa180-199: DVHRKENAGAAEKSITILST (SED ID  
NO:460)
- (5) aa320-339: LYNPERTITVKGNVETCAKA (SEQ ID  
NO:461)
- (6) aa340-359: EEEIMKKIRESYENDIASMN (SED ID  
NO:462)
- (7) aa370-389: LNALGLFPPTSGMPPTSGP (SEQ ID  
NO:463)
- (8) aa380-399: KIAPAEAPDAKVRMVIITGP (SED ID  
NO:464)

Out of these, six epitopes are common in both lung plural effusion fluid samples and in sera of lung patients. Of these six, SEQ ID NO:459 and SEQ ID NO:463 have no homology to other L523S-family proteins such as IGF-II mRNA-binding proteins 1 and 2. Accordingly, this indicates that these two peptides can be used as an assay system to determine the antibody response to L523S.

30

EXAMPLE 28

## GENERATION OF L523S-SPECIFIC CTL LINES USING IN VITRO WHOLE-GENE PRIMING

To determine if L523S is capable of generating a CD8<sup>+</sup> T cell immune response, CTLs were generated using *in vitro* whole-gene priming methodologies with tumor antigen-vaccinia infected DC (Yee et al, *The Journal of Immunology*, 157(9):4079-86, 1996), human CTL lines were derived that specifically recognize autologous fibroblasts transduced with the L523S tumor antigen, as determined by interferon-gamma ELISPOT analysis. Specifically, dendritic cells (DC) were differentiated from Percoll-purified monocytes derived from PBMC of normal human donors by plastic adherence and growing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following the five days of culture, the DC were infected overnight with a recombinant adenovirus that expresses L523S at a multiplicity of infection (M.O.I) of 33, 66 and 100, and matured overnight by the addition of 2 µg/ml CD40 ligand. The virus was then inactivated by UV irradiation. In order to generate a CTL line, autologous PBMC were isolated and CD8<sup>+</sup> T cells were enriched for by the negative selection using magnetic beads conjugated to CD4<sup>+</sup>, CD14<sup>+</sup>, CD16<sup>+</sup>, CD19<sup>+</sup>, CD34<sup>+</sup> and CD56<sup>+</sup> cells. CD8<sup>+</sup> T cells specific for L523S were established in round bottom 96-well plates using 10,000 L523S expressing DCs and 100,000 CD8<sup>+</sup> T cells per well in RPMI supplemented with 10% human serum, 10ng/ml of IL-6 and 5ng/ml of IL-12. The cultures were restimulated every 7-10 days using autologous primary fibroblasts retrovirally transduced with L523S, and the costimulatory molecule CD80 in the presence of IL-2. The cells were also stimulated with IFN-gamma to upregulate MHC Class I. The media was supplemented with 10U/ml of IL-2 at the time of stimulation as well as on days 2 and 5 following stimulation. Following three stimulation cycles, ten L523S specific CD8<sup>+</sup> T cell lines were identified using interferon-gamma ELISPOT analysis that specifically produce interferon-gamma when stimulated with the L523S tumor antigen-transduced autologous fibroblasts, but not with a control antigen.

One line, 6B1, was cloned using anti-CD3 and feeder cells. The clones were tested for specificity on L523S-transduced fibroblasts. In addition, using a panel of HLA-mismatched lines transduced with a vector expressing L523S and measuring interferon-gamma production by this CTL line in an ELISPOT assay, it was determined  
5 that this clone 6B1.4B8 is restricted by HLA-A0201.

Also using transfected Cos cells, it was shown that clone 6B1.4B8 recognizes Cos cells transfected with pcDNA3 HLA A0201/L523S in an HLA-restricted and antigen specific manner.

An epitope mapping study demonstrated the clone 6B1.4B8 recognizes  
10 HLA-A201 LCL loaded with peptide pool 3 (a polypeptide corresponding to amino acid positions 33-59 of L523S).

A peptide pool breakdown study demonstrated that clone 6B1.4B8 recognizes autologous B-LCL loaded with 15-mer peptides from amino acid positions 37-55 of L523S, TGYAFVCPDESWALKAIIE (SEQ ID NO:465). A further peptide  
15 breakdown study demonstrated that clone 6B1.4B8 recognizes T2 cells loaded with the same 15-mer peptides.

A peptide recognition study demonstrated that clone 6B1.4B8 prefers T2 cells loaded with the peptide FVDCPESWAL (SEQ ID NO:466) which corresponds to the amino acid sequence at positions 41-51 of L523S and is encoded by the DNA  
20 sequence of SEQ ID NO:467.

### EXAMPLE 29

#### L523S EXPRESSION IN OTHER HUMAN CANCERS

25 It was previously disclosed in Example 2 that L523S is expressed in lung cancers including squamous, adenocarcinoma and small cell carcinoma. EST profiling analysis of L523S further indicates that this protein may also be expressed in a number of other tumor types, including colon adenocarcinomas, prostate adenocarcinomas, CML, AML, Burkitt's Lymphoma, brain tumors, retinoblastomas, ovarian tumors,

teratocarcinomas, uterus myosarcomas, germ cell tumors as well as pancreatic and cervical tumor cell lines.

From the foregoing it will be appreciated that, although specific  
5   embodiments of the invention have been described herein for purposes of illustration,  
various modifications may be made without deviating from the spirit and scope of the  
invention. Accordingly, the invention is not limited except as by the appended claims.

## CLAIMS

1. An isolated polynucleotide comprising a sequence selected from the group consisting of:

(a) sequences provided in SEQ ID NOs:351, 353, 358, 362, 364, 366, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467;

(b) complements of the sequences provided in SEQ ID NOs:351, 353, 358, 362, 364, 366, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467;

(c) sequences consisting of at least 10 contiguous residues of a sequence provided in SEQ ID NOs:351, 353, 358, 362, 364, 366, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467;

(d) sequences that hybridize to a sequence provided in SEQ ID NOs:351, 353, 358, 362, 364, 366, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467, under highly stringent conditions;

(e) sequences having at least 75% identity to a sequence of SEQ ID NOs:351, 353, 358, 362, 364, 366, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467;

(f) sequences having at least 90% identity to a sequence of SEQ ID NOs:351, 353, 358, 362, 364, 366, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467; and

(g) degenerate variants of a sequence provided in SEQ ID NOs:351, 353, 358, 362, 364, 366, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467.

2. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) sequences having at least 90% identity to a polypeptide having an amino acid sequence of any one of the sequences provided in SEQ ID NOs:352, 354, 357, 361, 363, 365, 367, 369, 376-382, 387-419, 423, 427, 430, 433, 441, 443, 446, 449 and 451-466;



- (b) sequences encoded by a polynucleotide of claim 1;
- (c) sequences having at least 70% identity to a sequence encoded by a polynucleotide of claim 1; and
- (d) sequences having at least 90% identity to a sequence encoded by a polynucleotide of claim 1.

3. An expression vector comprising a polynucleotide of claim 1 operably linked to an expression control sequence.

4. A host cell transformed or transfected with an expression vector according to claim 3.

5. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a polypeptide of claim 2.

6. A method for detecting the presence of a cancer in a patient, comprising the steps of:

- (a) obtaining a biological sample from the patient;
- (b) contacting the biological sample with a binding agent that binds to a polypeptide of claim 2;
- (c) detecting in the sample an amount of polypeptide that binds to the binding agent; and
- (d) comparing the amount of polypeptide to a predetermined cut-off value and therefrom determining the presence of a cancer in the patient.

7. A fusion protein comprising at least one polypeptide according to claim 2.

8. A fusion protein according to claim 9, wherein the fusion protein is selected from the group consisting of sequences provided in SEQ ID NOs:352, 354, 423, 427, 430 and 433.

9. An oligonucleotide that hybridizes to a sequence recited in SEQ ID NOs:351, 353, 358, 362, 364, 366, 368, 370-375, 420, 424, 428, 431, 434, 442, 447, 450 and 467 under highly stringent conditions.

10. A method for stimulating and/or expanding T cells specific for a tumor protein, comprising contacting T cells with at least one component selected from the group consisting of:

- (a) polypeptides according to claim 2;
- (b) polynucleotides according to claim 1; and
- (c) antigen-presenting cells that express a polynucleotide according to claim 1,

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

11. An isolated T cell population, comprising T cells prepared according to the method of claim 10.

12. A composition comprising a first component selected from the group consisting of physiologically acceptable carriers and immunostimulants, and a second component selected from the group consisting of:

- (a) polypeptides according to claim 2;
- (b) polynucleotides according to claim 1;
- (c) antibodies according to claim 5;
- (d) fusion proteins according to claim 7;
- (e) T cell populations according to claim 11; and

(f) antigen presenting cells that express a polypeptide according to claim 2.

13. A method for stimulating an immune response in a patient, comprising administering to the patient a composition of claim 12.

14. A method for the treatment of a lung cancer in a patient, comprising administering to the patient a composition of claim 12.

15. A method for determining the presence of a cancer in a patient, comprising the steps of:

- (a) obtaining a biological sample from the patient;
- (b) contacting the biological sample with an oligonucleotide according to claim 9;
- (c) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and
- (d) compare the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence of the cancer in the patient.

16. A diagnostic kit comprising at least one oligonucleotide according to claim 9.

17. A diagnostic kit comprising at least one antibody according to claim 5 and a detection reagent, wherein the detection reagent comprises a reporter group.

18. A method for the treatment of lung cancer in a patient, comprising the steps of:

(a) incubating CD4+ and/or CD8+ T cells isolated from a patient with at least one component selected from the group consisting of: (i) polypeptides according to claim 2; (ii) polynucleotides according to claim 1; and (iii) antigen presenting cells that express a polypeptide of claim 2, such that T cell proliferate;

(b) administering to the patient an effective amount of the proliferated T cells,

and thereby inhibiting the development of a cancer in the patient.

## SEQUENCE LISTING

<110> Corixa Corporation  
 Wang, Tongtong  
 Marnerakis, Margarita  
 Fanger, Gary R.  
 Vedvick, Thomas S.  
 Carter, Darrick  
 Watanabe, Yoshihiro  
 Henderson, Robert A.  
 Peckham, David W.  
 Fanger, Neil

<120> COMPOSITIONS AND METHODS FOR THE THERAPY  
 AND DIAGNOSIS OF LUNG CANCER

<130> 210121.45502PC

<140> PCT

<141> 2001-06-28

<160> 467

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 315

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 236, 241

<223> n = A,T,C or G

<400> 1

```
gcagagacag actggtgggt gaacctggag gtgccaaaa agccagctgc gggcccagga 60
cagctgccgt gagactccc atgtcacagg cagtctgtgt gggtacagcg ccctcagtg 120
ttcatctcca gcagagacaa cggaggaggc tcccaccagg acggttctca ttatttatat 180
gttaatatgt ttgtaaactc atgtacagtt tttttgggg gggaagcaat ggaanggata 240
naaattacaa atagaatcat ttgctgtaat ccttaaatgg caaacgggtca ggccacgtga 300
aaaaaaaaaa aaaaaa 315
```

<210> 2

<211> 380

<212> DNA

<213> Homo sapiens

<400> 2

```
atttaggctt aagattttgt ttacccttgt tactaaggag caaattagta ttaaagtata 60
atatatataa acaaatataa aaagttttga gtggttcagc ttttttattt tttttaatgg 120
cataactttt aacaacactg ctctgtaatg ggttgaactg tggtaactcag actgagataa 180
ctgaaatcga tggatgtata gtgttattgc ataattatcc cactatgaag caaagggact 240
ggataaatcc ccagctctaga ttattagcct ttgttaacca tcaagcacct agaagaagaa 300
```

```
ttattggaaa tttgtcctc tgtaactggc actttggggg gtgacttatc ttttgccctt 360
gtaaaaaaaa aaaaaaaaaa 380
```

```
<210> 3
<211> 346
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<222> 316, 317, 318, 322, 323, 326, 329, 330, 331, 336, 337, 339,
340, 342, 343
<223> n = A,T,C or G
```

```
<400> 3
ttgtaagtat acaattttag aaaggattaa atgttattga tcattttact gaatactgca 60
catcctcacc atacaccatc cactttccaa taacatttaa tcctttctaa aattgtaagt 120
atacaattgt acctttcttg gattttcata acaaatatac catagactgt taattttatt 180
gaagtttctc taattggaat agtcattttt gtcttgtgct tttagaggta cctttgcttt 240
gacttccaac aatttgatca tatagtgttg agctgtggaa atctttaaagt ttattctata 300
gcaataaatt ctattnnnag annccngggn naaaannann annaaa 346
```

```
<210> 4
<211> 372
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<222> 297, 306, 332
<223> n = A,T,C or G
```

```
<400> 4
actagtctca ttactccaga attatgctct tgtacctgtg tggctggggt tcttagtcgt 60
tggttttggt tggttttttg aactggtatg taggggtggt caccgttcta atgtaagcac 120
tctcttctcc aagttgtgct ttgtggggac aatcattctt tgaacattag agaggaaggc 180
agttcaagct gttgaaaaga ctattgotta tttttgtttt taaagacctc cttgacgtca 240
tgtggacagt gcacgtgcot tacgctacat cttgttttct aggaagaagg ggaatgcnggg 300
aaggantggg tgctttgtga tggataaaac gnctaaataa cacaccttta cattttgaaa 360
aaaaaaaaac aa 372
```

```
<210> 5
<211> 698
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<222> 8, 345, 422, 430, 433, 436, 438, 472, 481, 486, 515, 521,
536, 549, 553, 556, 557, 559, 568, 593, 597, 605, 611, 613,
616, 618, 620, 628, 630, 632, 634, 635, 639, 643, 647, 648,
649, 652, 654, 658, 664, 690
<223> n = A,T,C or G
```

```
<400> 5
actagtanga tagaaacact gtgtcccgag agtaaggaga gaagctacta ttgattagag 60
cctaaaccag gttaactgca agaagaggcg ggatactttc agctttccat gtaactgtat 120
gcataaagcc aatgtagtcc agtttctaag atcatgttcc aagctaactg aatcccactt 180
```

```

caatacacac tcatgaactc ctgatggaac aataacaggc ccaagcctgt ggtatgatgt 240
gcacacttgc tagactcaga aaaaatacta ctctcataaa tgggtgggag tattttgggt 300
gacaacctac tttgcttggc tgagtgaagg aatgatattc atatnttcat ttatccaatg 360
gacatttagt tagtgctttt tatataccag gcatgatgct gagtgcacct cttgtgtata 420
tntccaaatn ttngtncngt cgctgcacat atctgaaatc ctatattaag antttcccaa 480
natgangtcc ctgggttttc caegccaact gatcngtcaa ngatctcacc tctgtntgtc 540
ctaaaaacgt ctncnntnang gttagacnng acctctcttc tccttcccg aanaatnaag 600
tgtngcaga nancncnncn ccccccctnnc tnncnctng cngctnnnc cncntgtngg 660
ggnggcgcgc ccgcggggg gaccccccn ttttccc 698

```

```

<210> 6
<211> 740
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 82, 406, 426, 434, 462, 536, 551, 558, 563, 567, 582, 584,
592, 638, 651, 660, 664, 673, 675, 697, 706, 711, 715, 716,
717, 723, 724, 725, 733
<223> n = A,T,C or G

```

```

<400> 6
actagtcaaa aatgctaaaa taatttggga gaaaatattt ttaaagtagt gttatagttt 60
catgtttatc ttttattatg tnttgtgaag ttgtgtcttt tcactaatta cctatactat 120
gccaatattt cttatatact atccataaca tttatactac atttgaaga gaatatgcac 180
gtgaaactta acactttata aggtaaaaat ggggtttcca agatttaata atctgatcaa 240
gttcttggta ttcccaataa gaatggactt ggtctgttaa ggggctaagg gagaagaaga 300
agataagggtt aaaaagtgtt aatgacaaaa cattctaaaa gaaatgcaaa aaaaaattta 360
tttccaagct ttcgaactat ttaaggaaag caaaatcatt tcctanatgc atatcatttg 420
tgagantttc toantaatat cctgaatcat tcatctcagc tnaggcttca tgttgactcg 480
atatgtcate tagggaaagt ctatttcatg gtccaaaact gttgccatag ttggttaggc 540
tttcttttaa ntgtgaanta ttacangaa attttctctt tnanagttct tnatagggtt 600
aggggtgtgg gaaaagcttc taacaatctg tagtgttnng tgttatctgt ncagaaccan 660
aatnacggaat cmanagaagg actgggtcta ttacangaa cgaatnatct ngtnnnnrtg 720
gtnnncaact cngggagacc 740

```

```

<210> 7
<211> 670
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 265, 268, 457, 470, 485, 546, 553, 566, 590, 596, 613, 624,
639, 653, 659, 661
<223> n = A,T,C or G

```

```

<400> 7
gctggggagc tcggcatggc ggtccccgct gcagccatgg ggccctcgcc gttgggcccag 60
agcggccccc gctcgatggc cccgtggtgc tcagtgcaga gcggcccgct gcgctactgt 120
cttgggatgc aggaactggt ccggggccac agcaagacgc cgagttcctg gcgcacagcg 180
ccaaggtgca ctcggtggcc tggagtgtgc acgggcctcg cctacctgag ggtcttcgac 240
aagacgcacg gtcttcttgc tgganaanga ccgttggtca aagaaaaaca ttatcgggga 300
catggggata gtgtggacca ctttgttggc atccaagtaa tcttgaccta tttgttacgg 360
cgtctggaga taaaaccatt cgcattctgg atgtgaggac tacaaaaatc attgccactg 420
tgaacactaa aggggagaac attaatatct gctggantcc tgatgggcan accattgtct 480
tagcnacaag gatgatgtgg tgactttatt gatgccaaag aaccocgttc caaagcaaaa 540

```

```

aaacantttc aanttcgaag tcaccnaaat ctctctggaac aatgaacatn aatatnttct 600
tcttgacaaat ggnccctggg tgnntcacat cctcagctnc cccaaaaactg aancctgtnc 660
natccacccc                                     670

```

```

<210> 8
<211> 689
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 253, 335, 410, 428, 448, 458, 466, 479, 480, 482, 483, 485,
488, 491, 492, 495, 499, 500, 502, 503, 512, 516, 524, 525,
526, 527, 530, 540, 546, 550, 581, 593, 594, 601, 606, 609,
610, 620, 621, 622, 628, 641, 646, 656, 673
<223> n = A,T,C or G

```

```

<400> 8
actagtatct aggaatgaac agtaaaagag gagcagttgg ctacttgatt acaacagagt 60
aaatgaagta ctggatttgg gaaaacctgg ttttattaga acatatggaa tgaagagcta 120
cacctagcat tgccctactta gccccctgaa ttaacagagc ccaattgaga caaacccctg 180
gcaacagggaa attcaaggga gaaaaagtaa gcaacttggg ctaggatgag ctgactccct 240
tagagcaaaag ganagacagc cccattacc aaataccatt ttgctctggg gcttgtgcag 300
ctggcagttgt tctgcccaca gcatggcacc ttatngtttt gatagcaact tcgttgaatt 360
ttcaccacact tattacttga aattataata tagcctgtcc gtttgcgtgn tccaggctgt 420
gatatatntt cctagtgggt tgacttttaa aataaaatnag gtttntttt cccccccnn 480
cnntnctncc nntnctcnnt cnntcccccc cncctngtcc tccnnntttn gggggggcnn 540
ccccnccggn ggacccccct ttggtccott agtggaggtt natggccctt ggnnttatcc 600
nggcctannn ttccccgtn nnaaatgntt cccctccaca ntccnccac ctcaanccgg 660
aagcctaagt ttntaccctg ggggtcccc                                     689

```

```

<210> 9
<211> 674
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 602, 632, 639, 668
<223> n = A,T,C or G

```

```

<400> 9
gtccactctc ctttgagtgt actgtcttac tgtgcactct gtttttcaac tttctagata 60
taaaaaatgc ttgttctata gtggagtaag agctcacaca ccaaggcgag caagataact 120
gaaaaaagcg aggttttttt gccaccttgg taaaggccag ttcactgcta tagaactgct 180
ataagcctga aggggaagttag ctatgagact ttccattttt cttagtcttc ccaataggct 240
ccttcattga aaaaggcttc ctgtaataat ttccacctaa tgaattagca gtgtgattat 300
ttctgaaata agagacaaat tggggcgcag agtcttctct tgatttaaaa taacaacccc 360
aaagttttgt ttggttttca ccaaaggaca tactctaggg ggtatgttgt tgaagacatt 420
caaaaacatt agctgttctg tctttcaatt tcaagttatt ttggagactg cctccatgtg 480
agttaattac tttgctctgg aactagcatt attgtcatta tcatcacatt ctgtcatcat 540
catctgaata atatttgtga ttccccctc tgcttgcato ttcttttgac tctctgggga 600
anaaatgtca aaaaaaaagg tcatctact cngcaaggnc catctaatca ctgcgctgga 660
aggaccnctc gcccc                                     674

```

```

<210> 10
<211> 346
<212> DNA

```



<213> Homo sapiens

<220>

<221> misc\_feature

<222> 320, 321, 322, 325, 326, 328, 329, 330, 332, 333, 334, 335, 342

<223> n = A,T,C or G

<400> 10

```
actagtctgc tgaatagaag cactatacat cctattgttt ctttctttcc aaaatcagcc 60
ttctgtctgt aacaaaaatg tactttatag agatggagga aaaggtctaa tactacatag 120
ccttaagtgt ttctgtcatt gttcaagtgt attttctgta acagaaacat atttggaaatg 180
ttttcttttt cccctataaa attgtaattc ctgaaatact gctgctttta aaagtccac 240
tgtcagatta tattatctaa caattgaata ttgtaaatat acttgcctta cctctcaata 300
aaaggggtact ttctatttan nnagnngnnn gnnnnataaa aaaaaa 346
```

<210> 11

<211> 602

<212> DNA

<213> Homo sapiens

<400> 11

```
actagtaaaa agcagcattg ccaataatc cctaatttcc cactaaaaat ataatgaat 60
gatgttaagc tttttgaaa gtttaggtta aacctactgt tgtagtagta atgtatttgt 120
tgcttccctt tatctggaat gtggcattag cttttttatt ttaacctct ttaattctta 180
ttcaattcca tgacttaagg ttggagagct aaacactggg atttttggat aacagactga 240
cagttttgca taattataat cggcattgta catgaaaagg atatggctac cttttgttaa 300
atctgcacct tctaaatata aaaaaaggga aatgaagtta taaatcaatt ttgtataat 360
ctgtttgaaa catgagtttt atttgcttaa tattagggct ttgccccttt tctgtaagtc 420
tctgggacat ctgtgtagaa ctgttctcat taaacaccaa acagttaagt ccattctctg 480
gtactagcta caaattcggg ttcattttct acttaacca ttaaataaac tgaatatatt 540
ctagatggtc tacttctgtt catataaaaa caaaacttga tttccaaaa aaaaaaaaaa 600
aa 602
```

<210> 12

<211> 685

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 170, 279, 318, 321, 322, 422, 450, 453, 459, 467, 468, 470, 473, 475, 482, 485, 486, 491, 498, 503, 506, 509, 522, 526, 527, 528, 538, 542, 544, 551, 567, 568, 569, 574, 576, 582, 587, 588, 589, 590, 592, 593, 598, 599, 603, 605, 608

<223> n = A,T,C or G

<221> misc\_feature

<222> 633, 634, 635, 644, 646, 648, 651, 655, 660, 662, 663, 672, 674, 675, 682, 683

<223> n = A,T,C or G

<400> 12

```
actagtctctg tgaaagtaca actgaaggca gaaagtgtta ggattttgca tctaagtgtc 60
attatcatgg tattgatgga cctaagaaaa taaaatttag actaagcccc caaataagct 120
gcatgcattt gtaacatgat tagtagattt gaatatatag atgtagattn ttgggtatct 180
aggtgtttta tcaattatga aaggaattaa agtaaaaggac ttgttagttg tttttattaa 240
atatgcatat agtagagtgc aaaaatatag caaaaatana aactaaaggt agaaaagcat 300
```

```

tttagatatg ccttaantnta nnaactgtgc cagtgggccc tcggaataga tgcaggcgag 360
agaccagtgc ctgggtgggtg cctccccctg tctgcccccc tgaagaactt cctcaccgtg 420
angtagtgga ctogtaggtg tcaactggan tantggganc agggccgnncn gtnanaagaa 480
ancanngtga nagtttncnc gtngangcng aactgtccct gngccnnnac gctcccanaa 540
cntntccaat ngacaatcga gtttccnnnc tccngnaacc tngccgnnnn cnngcccnnc 600
cantntgnta accccgcgcc cggatcgctc tcnntogtt ctcnncncaa ngggntttcn 660
cnnccgcgct cncnnccccc cnncc 685

```

&lt;210&gt; 13

&lt;211&gt; 694

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 503, 546, 599, 611, 636, 641, 643, 645, 656, 658, 662, 676, 679, 687

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 13

```

cactagtccac tcattagcgt ttccaatagg gctcttaagt ccagtagatt acgggtatgc 60
agttgacgaa gatctgggtt acaagaacta attaaatggt tcattgcatt ttgtgaagaa 120
cagaataatt ttataaaatg ttgttagttt ataattgcog aaaataattt aaagacactt 180
tttctctgtg tgtgcaaatg tgtgtttgtg atccattttt tttttttttt taggacacct 240
gtttactagc tagctttaca atatgccaaa aaaggatttc tccctgaccc catcogtggt 300
tcacccctctt ttccccccat gctttttgcc ctagtgttata acaaaggaat gatgatgatt 360
taaaaagtag ttctgtatct tcagtatctt ggtcttcacg aacctctcgg ttgggaaggg 420
gatcattttt tactgggtcat ttccctttgg agtgactac tttaacagat ggaaagaact 480
cattggccat ggaaacagcc gangtgttgg gagccagcag tgcattggcac cgtccggcat 540
ctggcctgat tggctcgtgt gccgtcattg tcagcacagt gccattggac atgggggaana 600
ctgactcgac ngccaatggt ttccatgaag aatacngcat ncnngtcat cactgnancc 660
angacgtat gggggncana gggccanttg ctcc 694

```

&lt;210&gt; 14

&lt;211&gt; 679

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 29, 68, 83, 87, 94, 104, 117, 142, 145, 151, 187, 201, 211, 226, 229, 239, 241, 245, 252, 255, 259, 303, 309, 359, 387, 400, 441, 446, 461, 492, 504, 505, 512, 525, 527, 533, 574, 592, 609, 610, 618, 620, 626, 627, 633, 639, 645, 654

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 14

```

cagccgcctg catctgtatc cagcgccang tcccgcagct ccagctgcg cgcgcccccc 60
agtcgccnac ccgttcggcc cangctnagt tagncctcac catnccggte aaaggangca 120
ccaagtgcac caaataactg cngtnccgat ntaaatctcat ctctctggctt cgcgggattg 180
ctgtccntgc cattggacta nggctccgat ncgactctca gaccanganc atcttcganc 240
naganactaa tnatnattnt tccagcttct acacaggagt ctatattctg atcggatccg 300
gcnccctcct gatgctgggt ggcttctga gctgctgcgg ggctgtgcaa gagtcccant 360
cagctgtggg actgttctct ggcttctctc tggatgatn cgccattgaa atacctgcg 420
ccatctgggg atattccact ncgatnatgt gattaaggaa ntccacggag ttttacaagg 480
acacgtacaa cnaactgaaa accnnnggat anccccaccg ggaanccnct aaagccatcc 540
actatgctt gaactgcaat gggttggctg gggnccttta acaatttaac cncatacatc 600
tggccccann aaaggacntn ctcgannccct tcnccgtgna attcngttct gatnccatca 660

```

cagaagtctc gaacaatcc

679

&lt;210&gt; 15

&lt;211&gt; 695

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

<222> 105, 172, 176, 179, 189, 203, 212, 219, 221, 229, 231, 238,  
242, 261, 266, 270, 278, 285, 286, 298, 311, 324, 337, 350,  
363, 384, 391, 395, 405, 411, 424, 427, 443, 448, 453, 455,  
458, 463, 467, 470, 479, 482, 484, 493, 499, 505, 518

&lt;223&gt; n = A,T,C or G

&lt;221&gt; misc\_feature

<222> 520, 523, 531, 540, 584, 595, 597, 609, 611, 626, 628, 651,  
652, 657, 661, 665, 669, 672, 681, 683, 691, 693

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 15

```
actagtggat aaaggccagg gatgctgctc aacctcctac catgtacagg gacgtctccc 60
cattacaact acccaatccg aagtgtcaac tgtgtcagga ctaanaaacc ctggttttga 120
ttaaaaaagg gcctgaaaaa aggggagcca caaatctgtc tgccttctca cnttantcnt 180
tggcaaatna gcattctgtc tcnttggctg cngcctcanc ncaaaaaanc ngaactcnat 240
cnggcccagg aatacatctc ncaatnaacn aaattganca aggcnnctggg aaatgccnga 300
tgggattatc ntccgcttgt tganccttcta agtttnttc ccttcattcn accctgccag 360
conagttctg ttagaaaaat gccngaatto naacnccggt ttontactc ngaattttaga 420
tctncanaaa ctctctggcc acnattcnaa ttanangmca ognacanatn ccttccatna 480
ancnaccocc acntttgana gccangacaa tgactgcntn aantgaagggc ntgaaggaa 540
aactttgaaa ggaaaaaaa cttttgttcc ggccctctcc aacnctctctg tgtttnacac 600
tgccctctng naaccttgga agcccnngna cagtggtaca tgttgttcta nnaaacngac 660
ncttnaatnt cnatcttccc nanaacgatt ncnc 695
```

&lt;210&gt; 16

&lt;211&gt; 699

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 299, 354, 483, 555, 571, 573, 577, 642, 651, 662, 667

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 16

```
gcgcgaagca gcagcgcagg ttgtcccggt ttccctccc ccttcccttc tccggttgcc 60
ttcccgggcc ccttacactc caccagtcocg gtcccgccat gtcccagaaa caagaagaag 120
agaacctctg ggaggagacc ggccgaggaga agcaggacac gcaggagaaa gaaggtattc 180
tgctcgtgag agctgaagag gcaaaagctaa aggccaaata cccaagccta ggacaaaagc 240
ctggaggctc cgacttcctc atgaagagac tccagaaaagg gcaaaagtc tttgactcng 300
gagactacaa catggccaaa gccaacatga agaataagca gctgccaaagt gcangaccag 360
acaagaacct ggtgactggt gatcacatcc ccaccccaca ggatctgcc agagaaagtc 420
ctcgctcgtc accagcaagc ttgcgggtgg ccaagttgaa tgatgctgcc ggggctctgc 480
canatctgag acgcttccct ccctgcccca cccgggtcct gtgctggctc ctgcccttcc 540
tgcttttgca gccangggct aggaagtggc ncnggtngtg gctggaaagc aaaacccttt 600
cctggttggt tccccccat ggagccctcg gggcgagccc angaacttga nctttttgt 660
tntcttccc 669
```

<210> 17  
 <211> 697  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 33, 48, 50, 55, 59, 60, 76, 77, 78, 90, 113, 118, 130, 135,  
 141, 143, 150, 156, 166, 167, 170, 172, 180, 181, 190, 192,  
 194, 199, 201, 209, 212, 224, 225, 226, 230, 233, 234, 236,  
 242, 244, 251, 253, 256, 268, 297, 305, 308, 311, 314  
 <223> n = A,T,C or G

<221> misc\_feature  
 <222> 315, 317, 322, 324, 327, 333, 337, 343, 362, 364, 367, 368,  
 373, 384, 388, 394, 406, 411, 413, 423, 429, 438, 449, 450,  
 473, 476, 479, 489, 491, 494, 499, 505, 507, 508, 522, 523,  
 527, 530, 533, 535, 538, 539, 545, 548, 550, 552, 555  
 <223> n = A,T,C or G

<221> misc\_feature  
 <222> 562, 563, 566, 568, 572, 577, 578, 580, 581, 591, 594, 622,  
 628, 632, 638, 642, 644, 653, 658, 662, 663, 665, 669, 675,  
 680, 686, 689  
 <223> n = A,T,C or G

<400> 17  
 gcaagatatg gacaactaag tgagaaggta atnctctact gctctagntn ctccngggcnn 60  
 gacgcgctga ggagannnac gctggcccan ctgccggcca cacacgggga tcntggtat 120  
 gcctgcccac gggancccca ncnctcggan ccatntcac acccgnnccn tncgccacn 180  
 ncnctggctcn cncngcccng nccagctcnc gnccccctcc gccnnnctcn tttnnctctc 240  
 cncnccctcc ncnacnaact cctaccncng gctccctccc cagcccccce ccgcaancct 300  
 ccacnacnc ntcnncnnga ancnccnctc gcnctcngcc cengcccctt gcccccgcgc 360  
 cncnacnncg cgnctcccgc cgcncgcngc ctncnccctt ccacnacag ncnacccgcg 420  
 agncaegcnc tccgcccncct gaogcccenn cccgcgcgcg tcaccttcat ggnccnncng 480  
 ccccgctcnc ncnctgcnc gccgcnngg cgcgccgcc cncnccngtn cncnccngng 540  
 cccngcngn angcngtgcg cncnccngcc gngccggnnc ncacccctcc ncnccgcgcc 600  
 cgcgcctcgg ggcctccgcg cncgcggntc antcccccnc cntnccgcc ctnctcngntc 660  
 cncnctcnc gctcngcgcn cgcncnccnc cccccc 697

<210> 18  
 <211> 670  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 234, 292, 329, 437, 458, 478, 487, 524, 542, 549, 550, 557,  
 576, 597, 603, 604, 646, 665  
 <223> n = A,T,C or G

<400> 18  
 ctgcgttgaa ggggtgcagta cctaagccgg agcggggtag aggcggggcg gcaccccctt 60  
 ctgacctcca gtgccgcggg cctcaagatc agacatggcc cagaacttga acgacttggc 120  
 gggacggctg cccgccgggc cccggggcat gggcacggcc ctgaagctgt tgctgggggc 180  
 cgcgcgcgtg gcctacgggtg tgccggaatc tgtgttcaac gtggaaggcg ggcncagagc 240  
 catctttctt aatcggtatg gtggagtga caggacacta tcctggggcg anggccttca 300  
 ctccaggatc cttggttcca gtaccccanc atctatgaca ttccggggcag acctcgaaaa 360

```

aatctcctcc  ctacaggctc  caaagaccta  cagatgggtg  atatctccct  gcgagtggtg  420
tctcgaccaa  tgctcangaa  cttcctaaca  tgttccancc  cctaagggct  ggactacnaa  480
gaacgantgt  tgccgtccat  tgtcacgaag  tgctcaagaa  ttnggtggc  caagttcaat  540
gmctcacnn  ctgatacccc  agcggggcca  agttancct  ggttgatccc  cgggganctg  600
acnnaaaag  gccaaaggact  tccccctcat  ctggataatg  tggcncctac  aaagctcaac  660
ttanccacc

```

```

<210> 19
<211> 606
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 506
<223> n = A,T,C or G

```

```

<400> 19
actagtgcac  acctcagctc  ccaggccagt  tctctgaatg  tcgaggagtt  ccaggatctc  60
tgccctcagt  tgcctctggt  tattgatggg  ggacaaattg  gggatggcca  gagccccgag  120
tgtgcctctg  gctcaactgt  ggttgatttg  tctgtgcccg  gaaagtttgg  catcattcgt  180
ccaggctgtg  cctcggaag  tactacagcc  atcctccaac  agaagtacgg  actgctcccc  240
tcacatcgct  cctacctgtg  aaactctggg  aagcaggaag  gcccaagacc  tgggtctgga  300
tactatgtgt  ctgtccactg  acgactgtca  aggcctcatt  tcgagaggcc  accggagcta  360
gggcactagc  ctgactttta  aggcagtgtg  tctttctgag  cactgtagac  caagcccttg  420
gagctgtctg  tttagccttg  cacctgggga  aaggatgtat  ttatttggat  ttccatatat  480
cagccaaaag  ctgaatggaa  aagtttagaa  cattctcagg  tggccttatt  ctaataaagt  540
tcttctgtct  gttttgtttt  tcaattgaaa  agttatataa  taacagattt  agaattcagt  600
gagacc

```

```

<210> 20
<211> 449
<212> DNA
<213> Homo sapiens

```

```

<400> 20
actagtaaac  aacagcagca  gaaacatcag  tatcagcagc  gtcgccagca  ggagaatatg  60
cagcgccaga  gccgaggaga  acccccgcct  cctgaggagg  acctgtccaa  actcttcaaa  120
ccaccacagc  cgccctgccg  gatggactcg  ctgctcattg  caggccagat  aaacacttac  180
tgccagaaac  tcaaggagtt  cactgccccaa  aacttaggca  agctcttcat  ggcccaggct  240
cttcaagaat  acaacaacta  agaaaaggaa  gtttccagaa  aagaagttaa  catgaactct  300
tgaagtcaac  ccaggccaac  tcttggaaga  aatatatttg  catattgaaa  agcacagagg  360
atttctttag  tgtcattgcc  gattttggct  ataacagtgt  ctttctagcc  ataataaaat  420
aaaacaaaat  cttgactgct  tgctcaaaa

```

```

<210> 21
<211> 409
<212> DNA
<213> Homo sapiens

```

```

<400> 21
tatcaatcaa  ctggtgaata  attaaacaat  gtgtggtgtg  atcatacaaa  gggtaccact  60
caatgataaa  aggaacaagc  tgccatatat  tggaacaaca  tggatgcatt  tcagaaactt  120
tatgttgagt  gaaagaacaa  acacggagaa  catactatgt  gtttctcttt  atgtaacatt  180
acagaaataa  aaacagaggc  aaccaccttt  gaggcagtat  ggagtggagt  agactggaaa  240
aagggaaggaa  ggaactctca  cgctgatgga  aatgtctgtg  tcttctattg  gtggtagtta  300
tgtggggata  tacatttgct  aaaattttat  gaactatata  ctaaaagaat  ctgcatttta  360
ttgggatgta  aataatacct  caattaaaaa  gacaaaaaaa  aaaaaaaaaa

```

<210> 22  
 <211> 649  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 263, 353, 610, 635, 646  
 <223> n = A,T,C or G

```
<400> 22
acaattttca ttatcttaag cacattgtac atttctacag aacctgtgat tattctcgca 60
tgataaggat ggtacttgca tatgttgaat tactactgtt gacagtttcc gcagaaatcc 120
tatttcagtg gaccaacatt gtggcatggc agcaaatgcc aacattttgt ggaatagcag 180
caaatctaca agagaccctg gttgggtttt cgttttgttt tctttgtttt ttcccccttc 240
tcttgaatca gcaggggatgg aangagggta ggggaagtat gaattaactcc ttccagtagt 300
agctctgaag tgtcacattt aatatcagtt ttttttaaac atgattctag ttnaatgtag 360
aagagagaag aaagagggaag tgttcacttt ttaatacac tgatttagaa atttgatgtc 420
ttatatcagt agttctgagg tattgatagc ttgctttatt tctgcttcta cgttgacagt 480
gttgaagcag ggtgaataac taggggcata tatatttttt tttttgttaa gctgtttcat 540
gatgttttct ttggaatttc cggataagtt caggaaaaca tctgcagtgt gttatctagt 600
ctgaagttcn tatcatctc attacaacaa aaacnccag aacgngntg 649
```

<210> 23  
 <211> 669  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 642, 661  
 <223> n = A,T,C or G

```
<400> 23
actagtgcgc tactggctga aatccctgca ggaccaggaa gagaaccagt tcagactttg 60
tactctcagt caccagctct ggaattagat aaattccttg aagatgtcag gaatgggata 120
tatctcttga cagcctttgg gctgctcggc cccagcagc caccagcagg ggaggtgaca 180
tcaactgtcg tgccccctc tgtcaagact ccgacacctg aaccagctga ggtggagact 240
cgcaaggtgg tgcctgatga gtgcaacatt gactcggtag agggaggagt caaacacac 300
ctgacacttc tgcgtgaagt ggaggacaaa ctgaacccgc acctgagctg tgacctgatg 360
ccaaatgaga atatccccga gttggcggct gagctggtag agctgggctt cattagttag 420
gctgaccaga gccggttgac ttctctgcta gaagagactt gaacaagttc aattttgcca 480
ggaacagtag cctcaactca gcgcctgtca ccgtctctcc ttagagctca ctccgggccag 540
gccctgatct gcgctgtggc tgtcctggac gtgctgcacc ctctgtcctt cccccagtc 600
agtattacct gtgaagccct tccctccttt attattcagg anggctgggg gggctccttg 660
nttctaacc 669
```

<210> 24  
 <211> 442  
 <212> DNA  
 <213> Homo sapiens

```
<400> 24
actagtacca tcttgacaga ggatacatgc tcccaaaacg ttgtttacca cacttaaaaa 60
tcactgccat cattaagcat cagtttcaaa attatagcca ttcattgatt actttttcca 120
gtgactatcc tcttttgat ttgtaagggt aaaaaaaca aaaaacaaaa 180
cttacgatgc acttttctcc agcacatcag atttcaaatt gaaaattaaa gacatgctat 240
```

```

ggtaatgcac ttgctagtag tacacacttt ggtacaacaa aaacagagg caagaaacaa 300
cggaagagaga aaagccttcc tttgttggcc cttaaactga gtcaagatct gaaatgtaga 360
gatgatctct gacgatacct gtatgttctt attgtgttaa taaaattgct ggtatgaaat 420
gacctaaaaa aaaaaaaga aa

```

```

<210> 25
<211> 656
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 330, 342, 418, 548, 579, 608
<223> n = A,T,C or G

```

```

<400> 25
tgcaagtacc acacactgtt tgaattttgc aaaaaaagtg actgtaggat caggtgatag 60
ccccggaatg tacagtgctt tgggtcacca agatgccttc taaaggctga cataccttgg 120
accctaattgg ggcagagagt atagccctag ccagtggtg acatgaccac tccctttggg 180
aggcctgagg tagaggggag ttggtatgtt tttctcagtg gaagcagcac atgagtgggt 240
gacaggatgt tagataaagg ctctagttag ggtgtcattg tcatttgaga gactgacaca 300
ctcctagcag ctggtaaagg ggtgctggan gccatggagg anctctagaa acattagcat 360
gggctgatct gattacttcc tggcatcccg ctcaacttta tgggaagtct tatttagang 420
atgggacagt tttccatctc cttgctgtgg agctctggaa cactctctaa atttccctct 480
attaaaaatc actgccctaa ctacacttcc tccttgaagg aatagaaatg gaactttctc 540
tgacatannt cttggcatgg ggagccagcc acaaatgana atctgaacgt gtccagggtt 600
ctcctganac tcactacat agaattgggt aaacctcccc ttggaataag gaaaaa 656

```

```

<210> 26
<211> 434
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 395
<223> n = A,T,C or G

```

```

<400> 26
actagttcag actgccacgc caaccccaga aaatacccca catgccagaa aagtgaagtc 60
ctaggtgttt ccatctatgt ttcaatctgt ccatctacca ggcttcgcga taaaaacaaa 120
acaaaaaaac gctgccaggt tttagaagca gttctggctc caaaaccatc aggatcctgc 180
caccagggtt cttttgaaat agtaccacat gtaaaaggga atttggcttt cacttcactc 240
aataactgaa ttgtcaggct ttgattgata attgtagaaa taagttagctc totgttggg 300
gaataagtta taatcagtat tcactctctt gttttttgtc actcttttct ctctaattgt 360
gtcatttgta ctgtttgaaa aatatttctt ctatnaaatt aaactaacct gccctaaaaa 420
aaaaaaaaaa aaaa

```

```

<210> 27
<211> 654
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 505, 533, 563, 592, 613, 635, 638
<223> n = A,T,C or G

```

```

<400> 27
actagtccaa cacagtcaga aacattgttt tgaatcctct gtaaaccaag gcattaatct 60
taataaaccac ggaatccattt aggtaccact tgatataaaa aggatatcca taatgaatat 120
tttatactgc atccctttaca tttagccacta aatacgtttat tgcttgatga agacctttca 180
cagaatccta tggattgcag catttcaactt ggctacttca taccatgcc ttaaagaggg 240
gcagttcttc aaaagcagaa acatgcgcgc agttctcaag ttttctcct aactcoattt 300
gaatgttaag gcagctggcc cccaatgtgg ggaggtccga acattttctg aattccatt 360
ttctgtttcg cggctaaatg acagtttctg tcattactta gattccgac ttcccaaaag 420
gtgttgattt acaagagggc cagctaatag cagaaatcat gaccttgaaa gttcagatga 480
attcaagctg tgagccaggg agganctcag tatggcaagg gtcttgagaa tcnngcattt 540
ggtacaaaaa aaatttttaa gcntttatgt tataccatgg aaccatagaa anggcaagg 600
aattgttaag aanaatttta agtgtccaga ccanaanga aaaaaaaaaa aaaa 654

```

```

<210> 28
<211> 670
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 101, 226, 274, 330, 385, 392, 397, 402, 452, 473, 476, 532,
534, 538, 550, 583, 595, 604, 613, 622, 643, 669
<223> n = A,T,C or G

```

```

<400> 28
ogtgtgcaca tactgggagg atttccacag ctgcacggtc acagccctta cggattgccca 60
ggagggggcg aaagatatgt gggataaact gagaaaagaa nccaaaaacc tcaaatccca 120
aggcagctta ttgcgaactct caacggggcg gcggggctcc tgctcccgcc 180
gttccggctg ctctctggtg ctctctcggc agcttttagc acctgncctt ccttctgagc 240
gtggggccag ctcccccgcc ggogcccacc cacnctcact coatgctccc ggaaatcgag 300
aggaagatca ttagtctttt ggggacgttn gtgattctct gtgatgtcga aaaaacatca 360
tatagggaat gtgggaaatc ctganctctt tnttatntct tntgatttct tgtgttttat 420
ttgcacaaat gttaccaatc agtgaccaac cnagcacagc caaaaatcgg aontcngctt 480
tagtccgtct tcacacacag aataagaaaa cggcaaaccc accccaactt tnanntttat 540
tattactaan ttttttctgt tgggcataag aatctcagga acngccctgg ggccnccgta 600
ctanagttaa ccnagctagt tncatgaaaa atgatgggct cncctcctaat gggaaagcca 660
agaaaaagnc
670

```

```

<210> 29
<211> 551
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 336, 474, 504, 511, 522, 523, 524, 540, 547
<223> n = A,T,C or G

```

```

<400> 29
actagtcttc cacagcctgt gaatccccct agacctttca agcatagatga goggagaaga 60
agatctcagc gtttagccac cttaccatcg cctgatgatt ctgtagaaaa ggtttcttct 120
ccctctccag ccaactgattg gaaagtattc tccatcagtt ctcaaaatca gcaagaatct 180
tcagtaccag aggtgcctga tgttgacat ttgccacttg agaagctggg accctgtctc 240
cctcttgact taagtctgtg ttcagaagtt acagcacggg tagcctcaga ttctctttac 300
cgtaatgaat gtcccagggc agaaaagag gatacnaga tgcctccaaa tctctcttcc 360
aaagcaatag ctgatgggaa gaggagctcc agcagcagca ggaatatcga aaacagaaaa 420
aaaagtgaaa ttgggaagac aaaagctcaa cagcattttg taaggagaaa aganaagatg 480
aggaagggaag agagaagaga gacnaagatc nctacggacc gnnncggaag aagaagaagn 540

```



aaaaaaaaa a

551

<210> 30  
 <211> 684  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 545, 570, 606, 657, 684  
 <223> n = A,T,C or G

```
<400> 30
actagtgtcta tctggaaaaa gccggggttg gaagaagctg tggagagtgc gtgtgcaatg 60
cgagactcat tctctggaaag catccctggc aaaaatgcag ctgagtacaa ggttatcact 120
gtgatagaac ctggactgct ttttgagata atagagatgc tgcagttctga agagacttcc 180
agcacctctc agttgaatga attaatgatg gcttctgagt caactttact ggctcaggaa 240
ccacgagaga tgactgcaga tgtaatcgag cttaaaggga aattcctcat caacttagaa 300
ggtggtgata ttctgtaaga gtcttctcat aaagtaattg tcatgccgac tacgaaagaa 360
aaatgccccc gttgttggaa gtatcacagc ggagtcttca gatacactgt gtccctcgatg 420
tgcagaagtt gtcagtgagg aaatagtatt aacagctcac tcgagcaaga accctcctga 480
cagtactggg ctagaagttt ggatggatta ttacaatat aggaagaaaa gccacaagaa 540
aggtnatgag tggatgagta aatggtggaan gatggggaat tcaaatcaga attatggaa 600
aagttnttcc tgttaactata gaaaggaatt atgttttatt acatgcagaa aatatana 660
tgtggtgtgt accgtggatg gaan 684
```

<210> 31  
 <211> 654  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 326, 582, 651  
 <223> n = A,T,C or G

```
<400> 31
ggcagaaaaa ggaaccaata ttccagaaac aagcttaata ggaacagctg cotgtacatc 60
aacatcttct cagaatgacc cagaagttat catcgtggga gctggcgtgc ttggctctgc 120
tttggcagct gtgctttcca gagatggaag aaaggtgaca gtcattgaga gagacttaaa 180
agagcctgac agaatagttg gagaatttct gcagccgggt ggttatacat ttctcaaa 240
ccttggtctt ggagatacac tggaaaggtc tgatgccag gttgtaaatg gttacatgat 300
tcatgatcag ggaagcaaaa tcagangttc agattcctta cctctgtgtc gaaaacaatc 360
aagtgcagag tggaaagagc ttccatcacg gaagattcat catgagcttc cggaaagcag 420
ctatggcaga gcccaatgca aagtttattg aaaggtgtgt gttacagtta ttagaggaag 480
atgatgtttg gatggagttt cagtacaagg ataaagagac tgggagatat caaggaaact 540
catgctccac tgactgttgt tcagatggg cttttctcca anttcaggaa aagcctggtc 600
tcaataaagt ttctgtatca ctcaatttgt tggcttctta tgaagaatgc nccc 654
```

<210> 32  
 <211> 673  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 376, 545, 627  
 <223> n = A,T,C or G

```

<400> 32
actagtgaag aaaaagaat tctgatacgg gacaaaaatg ctcttcaaaa catcatttctt 60
tatcacctga caccaggagt ttctatttga aaaggtattg aacctgggtg tactaacatt 120
ttaaagacca cacaaggagc caaaatcttt ctgaaagaag taaatgatac acttctgttg 180
aatgaattga aatcaaaaga atctgacatc atgacaacaa atggtgtaat tcatgttgta 240
gataaactcc tctatccagc agacacacct gttggaatg atcaactgct ggaataactt 300
ataaatttaa tcaaatcatc ccaaattaag ttgttcgtg gtagcacctt caagaaatc 360
cccgtagctg totatnagcc aattatttaa aaatacacca aactcattga tgggagtgcc 420
tgtgggaaat aactgaaaaa gagaccgaga agaacgaatc attacaggtc ctgaaataaa 480
atcacattga ttctactgag aggtggagaa acagaagaac tctgaagaaa ttgttacaag 540
aagangtccc aaggtcacca aattcattga aggtggtgat ggtctttatt tgaagatgaa 600
gaaattaaaa gacgcttcag ggagacnccc catgaaggaa ttgccagcca caaaaaatt 660
cagggattag aaa 673

```

```

<210> 33
<211> 673
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 325, 419, 452, 532, 538, 542, 571, 600, 616, 651, 653, 672
<223> n = A,T,C or G

```

```

<400> 33
actagtattt tactttcttc cgcttcagaa ggtttttcag actgagagcc taagcatact 60
ggatctgttg ttcttttttg gtctcacctc atcagtggtc atagtggcag aaattataaa 120
gaaggttgaa aggagcaggg aaaagatcca gaagcatggt agtctgcacat catcatcttt 180
tcttgaagta tgaatgcata tgcattattt tatttgcaaa cttaggaattg cagctctagg 240
atcatattaga agggcaagtt caagaggata tgaagatttg agaacttttt aactattcat 300
tgactaaaaa tgaacattaa tgttnaagac ttaagacttt aacctgctgg cagtcccaaa 360
tgaaattatg caactttgat atcatattcc ttgattttaa ttgggtcttt ttgatttgant 420
gaaactttat aaagcatatg gtcagttatt tnattaaaaa ggcaaaacct gaaccacctt 480
ctgcacttaa agaagtctaa cagtacaaat acctatctat cttagatgga tntatttntt 540
tntattttta aatattgtac tatttatggt nggtggggct ttctacttaa tacacaaatn 600
aattttatcat ttcaanggca ttctatttgg gtttagaagt tgattocaag nantgcatat 660
ttcgctactg tnt 673

```

```

<210> 34
<211> 684
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 414, 472, 480, 490, 503, 507, 508, 513, 523, 574, 575, 598,
659, 662, 675
<223> n = A,T,C or G

```

```

<400> 34
actagtttat tcaagaaaag aacttactga ttctctgtgt cctaaagcaa gagtggcagg 60
tgatcagggc tgggtgtaga tccggttctt ttagtgcagc taactgcatt tgtcactgat 120
gaccaaggag gaaatcacta agacatttga gaagcagtggt tatgaacggt ctgggacaag 180
ccacagttct gagccttaac cctgtagttt gcacacaaga acgagctcca cctccctctc 240
ttcaggagga atctgtgcgg atagattggc tggacttttc aatggttctg ggttgcaagt 300
gggcacttgt atggctgggt atggagcgga cagcccccagg aatcagagcc tcagcccggc 360
tgctcggttg gaaggtacag gtgttcagca ccttcggaaa aagggcataa agtngtgggg 420

```

```

gacaattctc agtccaagaa gaatgcattg accattgtgt gctatttgc tncctagtan 480
gaattggatn catttttgac cangatnntt ctncatgtct tntttgcaat gaaatcaaat 540
cccgattat ctacaagtgg tatgaagtc tgcnccccc agagaggctg ttcaggcnat 600
gtcttccaag ggcagggtgg gttacacat ttacctccc ctctccccc agattattgna 660
cncagaagga atttntttcc tccc                                     684

```

```

<210> 35
<211> 614
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 17, 20, 152, 223, 267, 287, 304, 306, 316, 319, 321, 355,
365, 382, 391, 407, 419, 428, 434, 464, 467, 477, 480, 495,
499, 505, 515, 516, 522, 524, 527, 542, 547, 549, 567, 572,
576, 578
<223> n = A,T,C or G

```

```

<400> 35
actagtccaa cgcgttngcn aatattcccc tggtagccta ctcccttacc ccgaatat 60
ggttaagatcg agcaatggct tcaggacatg ggttctcttc tctgtgtatc attcaagtgc 120
tcactgatg aagactggct tgtctcagtg tntcaacctc accagggctg tctcttggtc 180
cacacctcgc tcctgtttag tgccgtatga cagcccccac canatgacct tggccaaatc 240
acgggtttctc tgtggatcaat gttggtnggc tgattgggtg aaagtanggt ggaccaaagg 300
aagncncgtg agcagncanc nccagttctg caccagcagc gcctccgtcc tactnggggtg 360
tccngtttcc tcttgccctc gngtgggcta nggctgatt cggaanattg cctttgcang 420
gaagganga taantgggat ctaccaattg attctggcaa aacnatntct aagattnttn 480
tgctttatgt ggganacana tctantcttc atttntgtct gnanatnaca cctactcgt 540
gntcgancnc gtcttcgatt ttcgganaca cncantnaa tactggcgtt ctgtgtgtta 600
aaaaaaaaaa aaaa                                     614

```

```

<210> 36
<211> 686
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 222, 224, 237, 264, 285, 548, 551, 628, 643, 645, 665, 674
<223> n = A,T,C or G

```

```

<400> 36
gtggctggcc cggttctcgc attctcccca tccctacttt tctctccctc ctccctttcc 60
ctccctcgtc gactgtgtgc tgcgtgtcgc agactccctg accctccctc caccctcccc 120
taacctcggt gccaccggat tgcccttctt ttctgttgc ccagccacgc cctagtgtca 180
ggcggggggc cgtggagcgc cggaggcaact gcagcagaag anaaaaaga cagacnaac 240
ctcagctcgc cagtcccggtc gctngcttcc gcgcgcattg caatnagaca gacgccgctc 300
acctgctctg ggcacacgcg acccgtgggt gatttggcct tcagtggcat cacccttatg 360
ggtatttctt aatcagcgct tgcaaagatg gttaacctat gctacgcag ggagatacag 420
gagactggat tggaacattt ttgggttcta aaggtctgtt tggggtgcaa cactgaataa 480
ggatgccacc aaagcagccta cagcagctgc agatttcaca gcccaagtgt gggatgctgt 540
ctcagganat naattgataa cctggctcat aacacattgt caagaatgtg gatttcccca 600
ggatattatt atttgtttac cggggganag gataactgtt tcnctatttt taattgaaca 660
aactnaaaca aaanctaagg aatccc                                     686

```

```

<210> 37
<211> 681

```

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

<222> 7, 10, 11, 19, 25, 32, 46, 53, 77, 93, 101, 103, 109, 115,  
123, 128, 139, 157, 175, 180, 192, 193, 194, 212, 218, 226,  
227, 233, 240, 241, 259, 260, 267, 289, 296, 297, 298, 312,  
313, 314, 320, 325, 330, 337, 345, 346, 352, 353, 356  
<223> n = A,T,C or G

&lt;221&gt; misc\_feature

<222> 382, 385, 400, 427, 481, 484, 485, 491, 505, 515, 533, 542,  
544, 554, 557, 560, 561, 564, 575, 583, 589, 595, 607, 619,  
628, 634, 641, 645, 658, 670  
<223> n = A,T,C or G

&lt;400&gt; 37

gagacanaacn naacgtcang agaanaaaag angcatggaa cacaanccag gncgatggc 60  
caccttccca ccagcancca gcgccccca gcngcccca ngnccgang accangactc 120  
cancotgnat caatctganc tctattctcg gcccatncct acctcgagg tggangccgn 180  
aaaggtcgca cmmncagaga agctgctgcc anccacanc gccccnccc tgnccggctn 240  
nataggaaac tggtagcann gctgcanaat tcatacagga gcacgcgang ggacannnct 300  
cacactgagt tnnngatgan gcctnaccan ggacctnccc cagcnnaatg annacnngac 360  
tgccggaggaa ggaagacccc gnaacnggat ctggccggcn tgccaccccc ccaccctag 420  
gattatnccc cttgactgag tctctgagg gctaccgaa cccgcctcca ttccctacca 480  
natnntgctc natcgggact gacangctgg ggatnggagg ggctatcccc cancattcccc 540  
tnanaccaac agcnaacngan natnggggct cccngggctc ggngcaacnc tctncccc 600  
cggcgcnggc cttcgggtgt gtctctctcc aacnaattcc naaangcgcg gcccccngt 660  
ggactctctn ttgttccctc c 681

&lt;210&gt; 38

&lt;211&gt; 687

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

<222> 3, 30, 132, 151, 203, 226, 228, 233, 252, 264, 279, 306,  
308, 320, 340, 347, 380, 407, 429, 437, 440, 445, 448, 491,  
559, 567, 586, 589, 593, 596, 603, 605, 606, 609, 626, 639,  
655, 674, 682  
<223> n = A,T,C or G

&lt;400&gt; 38

canaaaaaaa aaaacatggc cgaiaaccagn aagctgcgcg atggcgccac ggcccctctt 60  
ctcccgcctc gtgtccggaa ggtttccctc cgaggcgccc cggctccgcg aagcggagga 120  
gagggcgagg cntgcccggg ccggagctca naggccctgg ggcgctctcg ctctcccgcc 180  
atcgcgaagg cggcgctaac ctanaggcctc cccgcaagg tcccnangc ggngcgccg 240  
gggggctgtg anaaccgcaa aaanaacgct gggcgcgcn gcaaccgcgc cacccccgcg 300  
aaggananac ttccacagan gcagcgcttc cacagcccan agccacnttt ctagggtgat 360  
gcaccccagt aagttcctgn cgggggaagt cacgcgtgtc aaaaaanctc ttgcctccac 420  
ggcgcgacna aggggangan ggcanagang tgcgcggccc acaggctatc tgatcacgtc 480  
gcccccctca nctcgtcttt gtgaattctc accttgttca accccacccg ccgtctctctc 540  
ctcctctgcg cttcctctna ccttaanaac cagcttctcc taccnnaatg tantnctctc 600  
gcnccngtng aaattaatc ggtccnccg aacctcttnc ctgtggcaac tgctnaaaga 660  
aactcgtgtt ctgnttactg cngtccc 687

```

<210> 39
<211> 695
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 300, 401, 423, 429, 431, 437, 443, 448, 454, 466, 492, 515,
523, 524, 536, 538, 541, 552, 561, 566, 581, 583, 619, 635,
636, 641, 649, 661, 694
<223> n = A,T,C or G

<400> 39
actagtctgg cctacaatag tgtgattcat gtaggacttc ttcatcaat tcaaaacccc 60
tagaaaaacg tatacagatt atataagtag ggataagatt tctaacattt ctgggctctc 120
tgacccctgc gctagactgt ggaaaggagg tattattata gtatacaaca ctgctgttgc 180
cttattagtt ataacatgat aggtgctgaa ttgtgattca caatttaaaa acactgtaat 240
ccaaactttt ttttttaact gtagatcatg catgtgaatg ttaatgttaa ttgttcaan 300
gttgttatgg gttagaaaaa ccacatgcct taaaatttta aaaagcaggg cccaaactta 360
ttagtttaaa attaggggta tgtttccagt ttgttattaa ntggttatag ctctgtttga 420
aanaaatcna ngaacangat ttngaaantt aagntgacat tattttccag tgacttgtta 480
atttgaaatc anacacggca ccttcctgtt tggtnctatt ggnntttgaa tccaancngg 540
ntccaaatct ntccggaac ngtcctttta acttttttac nanatottat ttttttattt 600
tggaatggcc ctatttaang ttaaaagggg ggggnnccac naccattent gaataaaact 660
naatatatat ccttgytccc ccaaatttta aggnng 695

<210> 40
<211> 674
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 403, 428, 432, 507, 530, 543, 580, 583, 591, 604, 608, 621,
624, 626, 639, 672
<223> n = A,T,C or G

<400> 40
aotagtagtc agttgggagt ggttgctata ccttgacttc atttatatga atttccactt 60
tattaaataa tagaaaaagaa aatcccgggtg ctgcagtag agttatagga cattctatgc 120
ttacagaaaa tatagccatg attgaaatca aatagtaaa gctgttctgg ctttttatct 180
tcttagctca tcttaataaa gtagtacact tgggatgcag tgctgtcgaa gtgctaatac 240
gttgtaacaa tagcaaaact cgaacttagg atgtgtttct tctctctctg gtttctgatt 300
tgatcaatto tttatttttg ggaacctata atacagtttt octattcttg gagataaaaa 360
ttaaatggat cactgatatt taagtcatc tgcttctcat ctnaatatct catattctgt 420
attagganaa antacctccc agcacagccc cctctcaaac cccacccaaa accaagcatt 480
tggaatgagt ctcttttatt tccgaantgt ggatgggata acccatatcn ctccaatttc 540
tgnttgggtt gggatttaatt ttgaactgtg catgaaaagn gnaaatcttt nctttgggtc 600
aaantttncg gtttaatttg nctngncaaa tccaatttnc ttttaagggtg tctttataaa 660
atttgcattt cngg 674

<210> 41
<211> 657
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

```

<222> 243, 247, 251, 261, 267, 272, 298, 312, 315, 421, 432, 434,  
501, 524, 569, 594, 607, 650

<223> n = A,T,C or G

<400> 41

```

gaaacatgca agtaccacac actgtttgaa ttttgacaaa aaagtgactg tagggatcag 60
gtgatagccc cggaatgtac agtgtcttgg tgcaccaaga tgccttctaa aggtcgacat 120
aocctggggc cotaatgggg cagagagtat agccctagcc cagtgggtac atgaccactc 180
cctttggggg gctgaagtta aagggaatgg tatgtgtttt ctcatggaag cagcacatga 240
atnggtacac ngatgttaaa ntaaggntct antttgggtg tcttgtcatt tgaaaaantg 300
acacactcct ancantctgt aaaggggtgc tgggaagccat ggaagaacct taaaaacatt 360
agcatgggct gatctgatta ctctcgtgca tcccgctcac ttttatggga agtcttatta 420
naaggatggg ananttttcc atatccttgc tgttggaaact ctggaacact ctctaaattt 480
ccctctatga aaaatcactg ncctactac acttctcctc tgaagggaata gaaatggacc 540
tttctctgac ttagtctctg gcatgggganc cagcccaaat taaaatctga ctntccgggt 600
ttctcngaa ctcactact tgaattggta aaacctcctt tggaattagn aaaaacc 657

```

<210> 42

<211> 389

<212> DNA

<213> Homo sapiens

<220>

<221> misc feature

<222> 179, 317, 320

<223> n = A,T,C or G

<400> 42

```

actagtctgt aggaatgtaa acaagtttgc tgggccttgc gagacttcac caggttgttt 60
cgatagctca cactcctgca ctgtgcctgt caccagcaggaa tgtctttttt aattagaaga 120
caggagaata acaaaaacca gactgtgtcc cacaatcaga aaacctccgtt gtggcagang 180
ggccttcacc gccaccaggg tgtcccgcca gacagggaga gactccagcc ttctgaggcc 240
atcctgaaga attcctgttt ggggggttgg aaggaaaaatc acccggattt aaaaagatgc 300
tgttgctcgc ccgcgtngtn ggggaaggac tggtttctct gtgaatttct taaaagaaaa 360
atattttaag ttaagaaaaa aaaaaaaaaa

```

<210> 43

<211> 279

<212> DNA

<213> Homo sapiens

<400> 43

```

actagtgaca agctcctggt cttgagatgt cttctcgtta aggagatggg ccttttggag 60
gtaaggata aatgaatgat gttctgtcat gattcactat tctagaacct gcatgacctt 120
tactgtgtta gctccttgaa tgttcttgaa attttagact ttctttgtaa acaataataa 180
tgtccttctc attgtataaa agctgttatg tgcaacagtg tggagatcct tgtctgattt 240
aataaaatac ttaaacactg aaaaaaaaaa aaaaaaaaaa

```

<210> 44

<211> 449

<212> DNA

<213> Homo sapiens

<220>

<221> misc feature

<222> 245, 256, 264, 266, 273, 281, 323, 325, 337, 393

<223> n = A,T,C or G

```

<400> 44
actagtagca tcttttctac aacgttaaaa ttgcagaagt agcttatcat taaaaaacia 60
caacaacaac aataacaata aatcctaagt gtaaatcagt tattctaccc cctaccaagg 120
atatcagcct gttttttccc tttttttccc ttgggaataat tgtgggcttc ttcccaaat 180
tctacagcct ctttctctct ctcacgtctg agcttccctg ttgcaacgca tgcgtttgctg 240
aagantgggc tgtttngctt ggantncgtt ccnagtggaa ncatgcttcc cctgtttact 300
gttggaaaga actcaaacct tcnancccta ggtgttncca ttttgtcaag tcatcaactgt 360
atttttctac tggcattaac aaaaaaagaa atnaaatatt gtccattaa actttaataa 420
aaccttaaaa gggaaaaaaa aaaaaaaaaa 449

```

```

<210> 45
<211> 559
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 263
<223> n = A,T,C or G

```

```

<400> 45
actagttggt gggaaatcac gacacttaaa gtcaatctgc gaaataattc ttttattaca 60
cactcaactga agttttttgag tcccagagag ccatctctatg tcaaacattc caagtactct 120
ttgagagccc agcattacat caacatgcco gtgcagttca aaccgaagtc cgcaggcaaa 180
tttgaagctt tgcttctcat tcaaacagat gaaggcaaga gtattgctat tcgactaatt 240
ggtgaagctc ttggaaaaaa ttactagaaa taacttttgg gtttaagttaa ttacataagt 300
tgtattttgt taacttttat tttctacact acaattatgc ttttgtatat atattttgta 360
tgatggatct ctataattgt agattttgtt tttacaagct aactactgaag actcgactga 420
aatattatgt atctagocca tagtattgta ctttaacttt acagggtgaa aaaaaaattc 480
tgtgtttgca ttgattatga tattctgaat aaatatggga atatatttta atgtgggttaa 540
aaaaaaaaa aaaaaggaa 559

```

```

<210> 46
<211> 731
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 270, 467, 477, 502, 635, 660, 671, 688, 695, 697, 725
<223> n = A,T,C or G

```

```

<400> 46
actagttota gtacatggc tgtcatagat gcaaccatta tattccattt agttttcttc 60
tcaggttccc taacaattgt ttgaaactga atatatatgt ttatgtatgt gttgtgtgtc 120
actgtcatgt atatgggtga tatgggatgt gtgcagtttt cagttatata tatattcata 180
tatacatatg catatatatg tataataatc atatatatcat gcatacactt gtataatata 240
catatatata cacatatatg cacacatatn atcaactgagt tccaaagtga gtcttttatt 300
ggggcaattg tattctctcc ctctgtctgc tcaactggcc ttgtcaagac atagcaattg 360
cttgatttcc ttggataag agtcttatct tcggcactct tgactctagc ctttaacttta 420
gattttctat ccagaataacc tctcatatct atcttaaaac ctaaganggg taagangtc 480
ataagattgt agtatgaag antttgctta gttaaattat atctcaggaa actcattcat 540
ctacaaatta aattgtaaaa tgatggtttg ttgtatctga aaaaatgttt agaacaagaa 600
atgtaacttg gtacactgta tatcaagaaa cctcnattta ttaagtctcc tcatagccan 660
atccttatat ngccctctct gacctgantt aatananact tgaataatga atagttaatt 720
taggnntggg c 731

```

```

<210> 47

```

<211> 640  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 5, 28, 106, 153, 158, 173, 176, 182, 189, 205, 210, 214,  
 225, 226, 229, 237, 260, 263, 269, 277, 281, 282, 322, 337,  
 338, 354, 365, 428, 441, 443, 456, 467, 476, 484, 503, 508,  
 554, 567, 575, 579, 588, 601, 606, 609, 611, 621, 636  
 <223> n = A,T,C or G

<400> 47  
 tgcgngcggg ttgtggcctt cttgttanga caotttcacg ogccctgaaa ttttcccgat 60  
 gtttaataac tcttcaggct cctgcctgca cagggttttt tcttantttg ttgcctaaca 120  
 gtacaccaaa tgtgacatcc tttcaccaat atngatttnc tcataccaca tontcnatgg 180  
 anacgactnc aacaattttt tgatnacccn aaanactggg ggctnnaana agtacctct 240  
 ggagcagcat ggacctgtcn gcnactaang gaacaanagt nntgaacatt tacacaacct 300  
 ttggtattgt ttactgttaa anagaaacat gcttctnncc ctgaccacag aggncaaccg 360  
 caganattgc caatgccaa gctgcagcgt tagatcaggt aatacattcc atggatgcat 420  
 tacatacatt gtccccgaaa nanaagatgc cctaanggct tcttcanact ggctcngaaa 480  
 aacnctacac ctggtgtgtg ganaacanaac tctttggaag atcatctggc acaagttccc 540  
 cccagtgggt tttnccttgg cactcattct accanactna ttcggaancc attctttggc 600  
 ntggcnttnt ntgtggacca ntcttctcac aactgnaccc 640

<210> 48  
 <211> 257  
 <212> DNA  
 <213> Homo sapiens

<400> 48  
 actagtatat gaaaatgtaa atatcacttg tgtactcaaa caaaagtgtg tcttaagctt 60  
 ccacottgag cagccttgga aacctaacct gctcttttta gcataatcac attttctaaa 120  
 tggattttct tgttcctgaa aaagtgtatt gtattagtgt tacatttgtt ttttgggaaga 180  
 ttatatattg atatgtatca tcataaaaata tttaataaaa aagtatcttt agagtgaaga 240  
 aaaaaaaaaa aaaaaaa 257

<210> 49  
 <211> 652  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 410, 428, 496, 571, 647  
 <223> n = A,T,C or G

<400> 49  
 actagttoag atgagtggct gctgaagggg ccccttctgc attttcatta taaccaaat 60  
 tccaattatt tgaactotta agtcatatat gtataatgac ttatgaatta gcacagttaa 120  
 gttgacacta gaaactgccc atttctgtat tacactatca aataggaaac attggaaaga 180  
 tggggaaaaa aatcttattt taaaatggct tagaaagttt tcagattact ttgaaaatcc 240  
 taaaactctt tctgtttcca aaacttgaaa atagttagat ggactcatgc attaagactg 300  
 ttttcaagc ttctctcaca tttttaaagt gtgattttcc ttttaataca catatttatt 360  
 ttctttaaag cagctatatc ccaacccatg acttttgaga tatacctatn aaaccaatat 420  
 aacagcagag ttattgaagc agctttctca aatgttgctt cagatgtgca agttgcaaat 480  
 tttattgtat tegtanaata caatttttgt tttaactgt atttcaatct atttctccaa 540  
 gatgttttct atatatagtg aaatatccca ngataactgc tctgtgtctg tcgcatttga 600



cgcataactg cacaatatgaa cagtgtatag ctcttggttg tgcattnacc cc 652

<210> 50

<211> 650

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 237, 270, 311, 443, 454, 488, 520, 535, 539, 556, 567, 594, 603, 634

<223> n = A,T,C or G

<400> 50

```

ttgcgcttg attttttttag ggcttgtgcc ctgtttcact tatagggtct agaatgcttg 60
tgttgagtaa aaaggagatg cccaatatgc aaagctgcta aatgtttctct ttgccataaa 120
gactccgtgt aactgtgtga acacttggga tttttctcct ctgtcccagag gtgcgtcgtct 180
gttttctttt ttgggttctt tctagaagat tgagaaatgc atatgacagg ctgagancac 240
ctccccaaac acacaagctc tcagccacan gcagcttctc cacagcccca gcttcgcaca 300
ggctcctgga nggctgcctg ggggaggcag acatgggagt gccagggtgg ccagatgggtt 360
ccaggactac aatgtcttta tttttaactg ttggccactg ctgccctcac coctgcccg 420
ctctggagta cgtctgccc canacaagtg ggantgaaat ggggttgggg gggaaacctg 480
attccanttt aggggggtgcc taactgaaca gtagggtatan aaggtgtgaa cctngnaant 540
gtttttataa attatntccc ttgttanatt tattttttaa ttaatatctc gttnaactgc 600
ccnnggaaaa ggggaaaaaa aaaaaaaat tctnttttaa cacatgaaca 650

```

<210> 51

<211> 545

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 66, 159, 195, 205, 214, 243, 278, 298, 306, 337, 366, 375, 382, 405, 446, 477, 492, 495, 503, 507, 508, 521, 537

<223> n = A,T,C or G

<400> 51

```

tggcgtgcaa ccagggtagc tgaagtttgg gtctgggact ggagattggc cattaggcct 60
cctganattc cagctccctt ccaccaagcc cagtcttgct acgtgggcaca gggcaaacct 120
gactcccttt gggcctcagt ttcccctccc ctctcatgana tgaaaagaat actacttttt 180
cttgtttgtc taacnttgct ggaacnaaag tgtngtcatt attgttgtat tgggtgatgt 240
gtncaaaaact gcagaagctc actgcctatg agaggaanta agagagatag tggatganag 300
ggacanaaag agtcattatt ttggtatgat ccaccctccc caacctttct ctccctcagt 360
cctgcncctc atgtntctgg tntgggtgat cctttgtgcc accanccatc atgcttttga 420
ttgtgcatat cctgggaagg ggggtgnatg tctcacaaact tgttgtcatc gtttgatanat 480
catgctttct tnatnaaaca aanaaannaa tgtttgacag ngttttaaata aaaaaaanaa 540
caaaa 545

```

<210> 52

<211> 678

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 98, 119, 121, 131, 136, 139, 140, 142, 143, 163, 168, 172, 176, 184, 189, 190, 191, 200, 201, 205, 207, 221, 223, 229,

230, 237, 240, 241, 255, 264, 266, 267, 276, 280, 288, 289,  
291, 297, 301, 306, 308, 314, 315, 326, 332, 335, 337

<223> n = A,T,C or G

<221> misc\_feature

<222> 339, 341, 343, 344, 345, 347, 350, 355, 356, 358, 362, 363,  
372, 379, 395, 397, 398, 400, 403, 412, 414, 421, 423, 431,  
435, 438, 439, 450, 457, 463, 467, 471, 474, 480, 483, 484,  
487, 490, 491, 492, 493, 499, 500, 504, 508, 518, 536

<223> n = A,T,C or G

<221> misc\_feature

<222> 538, 549, 551, 552, 554, 556, 557, 562, 563, 567, 571, 572,  
576, 579, 590, 592, 595, 598, 606, 609, 613, 620, 622, 624,  
626, 631, 634, 638, 641, 647, 654, 660, 661, 674

<223> n = A,T,C or G

<400> 52

actagttagaa gaacttttgcg gctttttgtgc ctctcacagg cgccataaagt catttgccatg 60  
ggaggaaagac gattttggggg gggaggggggg gggggcangg tccgtggggg ttccoctant 120  
ntatctccat ntccantggn cmtgtgcgc tcttccctcg tcnccatnga anttantccc 180  
tggnccecmn nocctctcmn nectnecct ccccccctcg ncnccctcmn cttttntan 240  
ncttcccat ctccctccc cctnanngtc ccacnccgn cagcaatnnc ncaactnctc 300  
nctccnccc tccnccgtt cttctttct cnaactntnc ncnntnccn tggcnntnaa 360  
annctctccc cmtgcaanc gattctctcc ctccnccnnc ctntccacte cntncttctc 420  
nncgctcct ntctctnnc ccacctctcn ccttcgnccc cantacnctc nccncccttn 480  
cgnntcttn nnntectcmn acnccccc tcccttncoc cctcttctcc cgggtntntc 540  
tctctccn ncnccnct cncnccctc nngcgnccnt ttccgcccen cncnccnt 600  
ccttctcm cantcatcn cntntccat nctnccctnc nctcacnccc gctnccccen 660  
ntctcttca caagctcc 678

<210> 53

<211> 502

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 139, 146, 215, 217, 257, 263, 289, 386, 420, 452, 457, 461,  
466, 482, 486

<223> n = A,T,C or G

<400> 53

tgaagatcct ggtgtgcgca tggggccgcg cccgcgccgt tgttaccggt attgtaagaa 60  
caagccgtac ccaaggtctc gcttctgcg aggtgtccct gatgccaaaa ttgcatttt 120  
tgacctgggg cggaataaang caaaantgga tgaagtctcg ctttggggc acatggtgtc 180  
agatcaatat gagcagctgt cctotgaagc cctgnangct gccgaattt gtgccaataa 240  
gtacatggtt aaaagntgt gnaagatgc ttccatatcc ggggtgcgnt ccaccccttc 300  
cacgtcatcc gcatcaacaa gatgttgtcc tgtgtgggg ctgacaggct cccaacaggc 360  
atgcgaagt cctttggaaa acccanggca ctgtggccag ggttcacatt gggccaattn 420  
atcatgttca tccgcaccaa ctgcagaaca angaactgt naattnaagc cctgccccag 480  
gnaanttca aatttcccg cc 502

<210> 54

<211> 494

<212> DNA

<213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 431, 442, 445  
 <223> n = A,T,C or G

<400> 54  
 actagtccaa gaaaaaatatg cttaatgtat attacaaagg ctttgtatat gttaacctgt 60  
 tttaatgccaa aaagtittgct ttgtccacaa ttctcttaag acctcttcag aaagggattt 120  
 gtttgccctta atgaatactg ttgggaaaaa acacagtata atgagtgaaa agggcagaag 180  
 caagaatttt ctacatctta ggcactccaa gaagaatgag tatccacatt tagatggcac 240  
 attatgagga ctttaaatctt tccttaaaaca caataatggt ttcttttttc ttttattcac 300  
 atgattttcta agtatatttt toatgcagga cagtttttca acctgtatgt acagtgcactg 360  
 tgttaaattt ttctttcagt ggcaacctct ataactctta aaatatgggt agcatcttgt 420  
 ctgtttttgaa ngggatatga cnatnaatct atcagatggg aaatcctggt tccaagttag 480  
 aaaaaaaaaa aaaa 494

<210> 55  
 <211> 606  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 375, 395, 511, 542, 559, 569, 578, 581  
 <223> n = A,T,C or G

<400> 55  
 actagtaaaa agcagcattg ccaaataatc cctaattttc cactaaaaat ataatgaaat 60  
 gatgttaagc tttttgaaaa gtttaggtta aacctactgt tgttagatta atgtatttgt 120  
 tgcttccott tatctggaat gtggcattag ctcttttatt ttaacctct ttaattctta 180  
 ttcaattcca tgacttaagg ttggagagct aaacactggg atttttggat aacagactga 240  
 cagttttgca taattataat cggcattgta catagaaaag atatggctac cttttgttaa 300  
 atctgcactt tctaaatatt aaaaaaggga aatgaagtat aaatcaattt ttgtataatc 360  
 tgtttgaaac atganttta ttgtcttaat attanggctt tgccttttc ttgtagtctc 420  
 ttgggtacct gtgtaaaact gttctcatta aacaccaaac agttaagtcc attctctggt 480  
 actagctaca aattccgttt catattctac ntaacaattt aaattaactg aaatatttct 540  
 anatggtcta cttctgtctn taataaaacna aacttgantt nccaaaaaaa aaaaaaaaaa 600  
 aaaaaa 606

<210> 56  
 <211> 183  
 <212> DNA  
 <213> Homo sapiens

<400> 56  
 actagtatat ttaaaacttac aggcttattt gtaatgtaaa ccaccatttt aatgtactgt 60  
 aattaacatg gttataatac gtacaatctt tcctcatccc catcacacaa ctttttttgt 120  
 gtgtgataaa ctgatttttg ttgtcaataa aaccttgaaa aataaaaaaa aaaaaaaaaa 180  
 aaa 183

<210> 57  
 <211> 622  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 358, 368, 412, 414, 425, 430, 453, 455, 469, 475, 495, 499,

529, 540, 564, 575, 590

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 57

```

actagtcaact actgtcttct ccttgtagct aatcaatcaa tattcttccc ttgcctgtgg 60
gcagtggaga gtgctgctgg gtgtacgtg cacctgcca ctgagttggg gaaagaggat 120
aatcagtgag cactgtttctg ctacagagctc ctgatctacc ccacccccta ggatccaggaa 180
ctgggtcaaa gctgcattgaa accaggccct ggcagcaacc tgggaatggc tggaggtggg 240
agagaacctg actctctctt cctctccct cctccaacat tactggaact ctatcctgtt 300
agggatcttc tgagcttggt tccctgctgg gtgggacaga agacaaagga gaaggggagg 360
tctacaanaa gcagcccttc tttgtctctt ggggttaatg agcttgacct anantcoatg 420
gaganaccan aagcctctga tttttaattt ccntnaaatg tttgaagnt atantacat 480
atatatat tttttnaant ttgagctctt gatattgtctt aaaatccant cctctgtgcn 540
gaaacttgaa ttaaaacat gaanaaaat gtttncctta aagatgttan taattaattg 600
aaacttgaaa aaaaaaaaaa aa

```

622

&lt;210&gt; 58

&lt;211&gt; 433

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 58

```

gaacaaattc tgatttggtta tgtaccgtca aaagacttga agaaatttca tgattttgca 60
tgttggaagc gttgaaaatt gaaagttact gcttttccac ttgctcatat agtaaaagga 120
tcctttcagc tgccagtggt gaataatgta tcatccagag tgatgttacc tgtgacagtc 180
accagcttta agctgaacca ttttatgaat accaaataaa tagacctctt gtaactgaaa 240
catatttggt actttaatcg tgcgtgctgg atagaaatat ttttactggt tttctctgaat 300
tgacagtaaa cctgtccatt atgaatggcc tactgttcta ttatttggtt tgacttgaat 360
ttatccacca aagacttcat ttgtgtatca toaataaagt ttgtatgttc aactgaaaaa 420
aaaaaaaaaa aaa

```

433

&lt;210&gt; 59

&lt;211&gt; 649

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 22, 190, 217, 430, 433, 484, 544, 550, 577, 583, 594

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 59

```

actagttatt atctgacttt cnggttataa tcattctaag gagtgtgaag tagcctctgg 60
tgtcatttgg atttgcattt ctctgatgag tgatgctatc aagcaccttt gctgggtgctg 120
ttggccatat gtgtatgttc cctggagaag tgtctgtgct gagccttggc ccaacttttta 180
attaggcgtn tgtcttttta ttactgagtt gtaaganttc tttatatatt ctggatttcta 240
gaacctatc agatacatgg tttgcaaata ttttctccca ttctgtgggt tgtgttttca 300
ctttatcgat aatgtcctta gacataatat aaatttggat tttaaaagtg acttgatttg 360
ggctgtgcaa ggtgggtcaa cgcttgtaat cccagcactt tgggagactg aggtgggtgg 420
atcatatgan gangctagga gttcgaggtc agcctggcca gcataggcaa aacttgtctc 480
tacnaaaat acaaaaatta gtcaggcatg gtggtgcacg tctgtaatac cagcttctca 540
ggangctgan gcacaaggat caactgaacc ccagaangaa gangttgcag tganctgaag 600
atcatgccag ggcaacaaaa atgagaactt gtttaaaaaa aaaaaaaaaa

```

649

&lt;210&gt; 60

&lt;211&gt; 423

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 209, 222, 277, 389, 398  
 <223> n = A,T,C or G

<400> 60  
 actagttoag gcoctccagtc tcaactgacaa acatggggaa gtgtgcccag ctggctggaa 60  
 acctggcagc gataccatca agcctgatgt ccaaaagagc aaagaatatt tctccaaaga 120  
 gaagtgcagc ctgggctggt ttagtgccag gctgcgggtg gcagccatga gaacaaaacc 180  
 tctctctgat ttttttttcc cattagatana acacaagact cngattcagc cgaattgtgg 240  
 tgtcttacaa ggcagggtct tctacacagg ggtgganaaa acagccttcc ttccttttgt 300  
 aggaatggcc tgagtggcgc ttgtgggcag gctaactggt tgtatgatgt attagtagag 360  
 caaccocatta atcttttgtta gtttgtatna aacttganct gagaccttaa acaaaaaaaa 420  
 aaa 423

<210> 61  
 <211> 423  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 195, 285, 295, 329, 335, 340, 347, 367, 382, 383, 391, 396,  
 418  
 <223> n = A,T,C or G

<400> 61  
 cgggactgga atgtaaagtg aagttcggag ctctgagcac gggtctctcc cgccgggtcc 60  
 tccctcccca gaccocagag ggagaggcag accccgccca gccccgcctc agccctgct 120  
 caggtcttag tatggctggg agtcgggggc cacaggcctc tagctgtgct gctcaagaag 180  
 actggatcag ggtanctaca agtggccggg ccttgccctt gggattctac cctgttctca 240  
 atttgggtgt ggggtgcggg gtcctctggc cccctttcca cactnccctc ctcengacag 300  
 caacotccct tggggcaatt gggcctggnt ctcncccggt tgttgcnacc ctttgttggt 360  
 ttaaggngctt taaaaatggt annttttccc ntgcnggggt taaaaaagga aaaaactnaa 420  
 aaa 423

<210> 62  
 <211> 683  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 218, 291, 305, 411, 416, 441, 443, 453, 522, 523, 536, 542,  
 547, 566, 588, 592, 595, 603, 621, 628, 630, 632, 644, 645,  
 648, 655, 660, 672, 674, 676, 677, 683  
 <223> n = A,T,C or G

<400> 62  
 gctggagagg ggtacggact ttcttgagg tgtccagggt tggaaatgaga ctgaactcaa 60  
 gaagagacc taagagactg gggaatgggt cctgccttca ggaaagtga agacgcttag 120  
 gctgtcaaca cttaaaggaa gtcccttga agcccgagct ggacagacta gaccattga 180  
 tggggccact ggccatggct cgtggacaag acattcngt gggccatggc acacggggg 240  
 ggatcaaaat gtgtacttgt ggggtctcgc ccttgccaa aacaaaacca ntcccactcc 300  
 tgtcnttga ctttcttccc attccctcct ccccaaatgc acttcccctc ctcctctcgc 360  
 cctcactgtg tttttggaat tctgtttccc tcaaaattgt taatttttta nttingaacc 420  
 atgaacttat gtttggggtc nangttcccc ttccaatgc atactaatat attaatggtt 480

```

attttttttt gaaatatttt ttaatgaact tggaaaaaat tnnatgaatt tcttntcttc 540
cnnntttttt ggggggggtg gggggntggg ttaaaatttt ttggaancc cnatnggaaa 600
tnttaacttg gggcccccct naaaaaantn anttccaatt cttnnatngc cccntntccn 660
ctaaaaaaa ananannaaa aan
683

```

```

<210> 63
<211> 731
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 237, 249, 263, 288, 312, 317, 323, 326, 337, 352, 362, 370,
377, 400, 411, 414, 434, 436, 446, 457, 473, 486, 497, 498,
502, 512, 531, 546, 554, 563, 565, 566, 588, 597, 608, 611,
613, 615, 627, 632, 640, 641, 644, 654, 660, 663, 665
<223> n = A,T,C or G

```

```

<221> misc_feature
<222> 671, 678, 692, 697, 698, 699, 704, 705, 712, 714, 717, 718,
719, 723, 725, 730, 731
<223> n = A,T,C or G

```

```

<400> 63
actagtcata aaggggtgtgc gcgtcttcga cgtggcggtc ttggcgcac tgctgcgaga 60
cccgccctgt gacctcaagg tcatccactt ggtgcgtgat ccccgccggtg ttggcgagttc 120
aaggatccgc tcgcgcacac gccctcatcg tgagagcccta cagggtggtgc gcagccgaga 180
ccgcgagctc accgcattgc cttcttggag ccgcgggggc acaagcttgg cgccanaaa 240
gaaggcgctn ggggcccgcga aantaccacg cctctggggc tatggaangt cctcttgcga 300
taatatgggt tnaaaanctg canaanagcc cctgcancoc cctgaactgg gntgcagggc 360
cncttacctn gtttggntgc ggttacaagg aacctgtttn ggaaaaacct ncnnaaaacc 420
ttccgggaaa attntncaa ttttntttgg ggaattnttg ggtaaaaccc cnnaaaatgg 480
gaaacntttt tgcctnnaa antaaaccat tnggttcggt gggccccccc ncaaaacctt 540
ttttnttttt tttntgcccc cantnncccc ccggggcccc ttttttntng ggaaaaancc 600
ccccctncc nanantttta aaaggngngg anaatttttn nttncccccc gggcccccn 660
gngngtaaaa nggttttenc cccccgaggg gngggggnnc ctnnaaaacc cntntcnna 720
cnonttttn n
731

```

```

<210> 64
<211> 313
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 240
<223> n = A,T,C or G

```

```

<400> 64
actagtgtgt caaaccacga ctgaagaaag acgaaaagtg ggaaataact tgcaacgtct 60
gttagagatg gttgctacac atgttgggtc ttagagaaaa catcttgagg agcagattgc 120
taaaagtgtat agagaatatg aagaatgcac gtcagaagat ctctcggaata atattaaaga 180
gattagagat aagtatgaga agaaagctac tctaattaa gttctctgaag aatgaagatn 240
aaatgttgat catgtatata tatccatagt gaataaaatt gtctcagtaa agttgtaaaa 300
aaaaaaaaaaa aaa
313

```

```

<210> 65
<211> 420

```

```

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 400, 402, 403, 404, 405, 406, 409, 411, 412, 414, 415, 416
<223> n = A,T,C or G

<400> 65
actagtctcc tggcaggcaa gggcttccaa ctgaggcagt gcatgtgtgg cagagagagg 60
caggaagctg gcagtggcag ctctctgtgc tagggagggg tgtgggtccc tcctccctg 120
tctgggaggt tggagggaa aatctaggcc ttagcttgcc ctctgccac ccttcccctt 180
gtagatactg ccttaacaact cctcctctc tcagctgtgg ctgccaccca agccaggttt 240
ctccgtgctc actaatttat ttccaggaaa ggtgtgtgga agacatgagc cgtgtataat 300
atttgtttta acattttcat tgcaagtatt gaccatcatc cttggttgtg tatcgttgta 360
acacaaatta atgatattaa aaagcatcca aacaaagccn annnnaana nnnnnngaaa 420

<210> 66
<211> 676
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 328, 454, 505, 555, 586, 612, 636, 641
<223> n = A,T,C or G

<400> 66
actagtcttc tatgatcatt aaatcattc tcagggttaa gaaaggaatg taaattttgt 60
ctcaaatttg taattcatca ataagttttt gaagagtcca gatttttagt caggtcttaa 120
aaataaaact acaaatcttg atgcattttt aaattctgca aatgtttcct ggggtgactt 180
aacaaggaaat aatcccacaa tatacctagc taactaatac atggagctgg ggctcaaccc 240
actgttttta aggatttgcg ctacttgtg gctgaggaaa aataagtagt tcaggaggaa 300
gtagttttta aatgtgagct tatagatngg aaacagaata tcaacttaat tatggaaatt 360
gttagaaacc tgttctcttg ttatctgaat cttgattgca attactattg tactggaatg 420
actccagccc attgcaaaagt ctacagatc ttanctgtgt agttgaattc cttggaaatt 480
ctttttaaga aaaaattgga gtttnaaaga aataaacccc tttgttaaat gaagcttgcc 540
tttttggtga aanaaatca tccgcaggg cttattgttt aaaaanggaa ttttaagcct 600
ccctggaaaa anttgttaat taaatgggga aaatgntggg naaaaattat ccgttagggt 660
ttaaagggaa aactta 676

<210> 67
<211> 620
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 419, 493, 519, 568, 605, 610
<223> n = A,T,C or G

<400> 67
caccattaaa gctgcttacc aagaacttcc ccagcatttt gacttccctg ttgatagct 60
gaattgtgag caggtgatag aagagccttt ctagtgtaac atacagataa ttgtgtgaat 120
acattccatt taatgaaggg gttacatctg ttacgaagct actaagaagc agcaagagca 180
taggggaaaa aaatctgctc agaacgcac aaactcacat gtgcccctc tactacaaac 240
agattgtagt gctgtggtgg tttattccgt tgtgcagaac ttgcaagctg agtcactaaa 300

```

```

cccaagaga ggaattata ggttagtta acattgtaac ccaggaaact aagtttaatt 360
cacttttgaa gtgttttgtt ttttattttt ggtttgtctg atttactttg ggggaaaaang 420
ctaaaaaaa agggatatca atctctaatt cagtgcaccac taaaagtgtg ccoctaaaaag 480
tctttactgg aanttatggg actttttaag ctccaggntt ttgtgtctcc caaattaacc 540
ttgcatgggc cccctaaaaa tgttgaangg cattctctgcc tctaagtttg ggggaaaattc 600
cccntttttt aaaatttga 620

```

<210> 68

<211> 551

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 286, 464, 480, 501, 502, 518, 528, 533, 536, 537, 538, 539, 540, 541, 543, 544, 545, 547, 548, 549

<223> n = A,T,C or G

<400> 68

```

actagtagct ggtacataat cactgaggag ctattttctta acatgctttt atagaccatg 60
ctaagtctag accagttatt aagggtctaat ctccaccctc cttagctgta agagtctggc 120
tagaacaga cctctctgtg caataacttg tggccactgg aaatccctgg gccggcattt 180
gtattggggt tgcaatgact cccaagggcc aaaagagtta aaggcaagac tgggatttct 240
ctgagactgt tgggtgaaact cctccaagg ctgagggggt cagtangtgc tctggggagg 300
actcggcacc actttgtat tcaacaagcc acttgaagcc caattataaa attgttattt 360
tacagctgat ggaactcaat ttgaaccttc aaaactttgt tagtttatcc tattatatgt 420
ttaaaccotaa ttacatttgt ctagcatttg atttgggttc tgtngcatat gtttttttcc 480
cctatgtgct cccctcccc nnatcttaac ttaaaacnca attttgcnat tcnccnnnnn 540
nannnnanna a 551

```

<210> 69

<211> 396

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 235, 310, 323, 381

<223> n = A,T,C or G

<400> 69

```

cagaaatgga aagcagagtt ttcattttctg tttataaacg tctccaaaca aaaaaggaaa 60
gcagagtttt cattaaatcc ttttaacctt tttttttctt ggtaatcccc tcaataaaca 120
gtatgtggga tattgaatgt taaggaggata tttttttcta ttatttttat aattgtacaa 180
aattaagcaa atgttaaaag ttttatatgc tttattaatg ttttcaaaag gtatnataca 240
tgtgatacat tttttaagct tcagttgctt gtctcttggt actttctgtt atggggtttt 300
ggggagccan aaaccaatct acnatctctt ttgttttgcc aggacatgca ataaaattta 360
aaaaataaat aaaaactatt nagaatttga aaaaaa 396

```

<210> 70

<211> 536

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 388, 446, 455

<223> n = A,T,C or G



&lt;400&gt; 70

```

actagtgcgaa aagcaaatat aaacatcgaa aaggcggtcc tcacgttagc tgaagatata 60
cttcgaaaga cccctgtaaa agagcccaac agtgaataatg tagatatcag cagtgaggga 120
ggcgtagacag gctggaagag caaatgctgc tgagcattct cctgttccat cagttgccat 180
ccactacccc gttttctctt ctgtctgcaa aataaaccac tctgtccatt ttttaactta 240
aacagatatt tttgtttctc atcttaacta tccaagccac ctattttatt tgttctttca 300
tctgtgactg ctgtgtgact ttatcataat ttcttcaaaa caaaaaaatg tatagaaaaa 360
tcattgtctgt gacttcattt ttaaatgnta cttgtctcag tcaactgcgt ttcagttggt 420
ttatagtcca gttcttatca acattnaaac ctatngcaat catttcaaat ctattctgca 480
aattgtataa gaataaaagt tagaatttaa caattaaaaa aaaaaaaa 536

```

&lt;210&gt; 71

&lt;211&gt; 865

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

```

<222> 22, 35, 39, 56, 131, 138, 146, 183, 194, 197, 238, 269, 277,
282, 297, 316, 331, 336, 340, 341, 346, 349, 370, 376, 381,
382, 392, 396, 397, 401, 433, 444, 445, 454, 455, 469, 472,
477, 480, 482, 489, 497, 499, 511, 522, 526, 527

```

&lt;223&gt; n = A,T,C or G

&lt;221&gt; misc\_feature

```

<222> 545, 553, 556, 567, 574, 580, 610, 613, 634, 638, 639, 663,
672, 689, 693, 694, 701, 704, 713, 723, 729, 732, 743, 744,
749, 761, 765, 767, 769, 772, 774, 780, 783, 788, 792, 803,
810, 824, 840, 848

```

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 71

```

gacaaagcgt taggagaaga anagaggcag ggaanactnc ccagggcaga tggcncctt 60
cccaccagca accagcgccc ccaccagccc ccagggcccg gacgacgaag actccatcct 120
ggattaatct nacctctntc gctgncacca ttctactctc ggaggtggag gccggaaaag 180
tcncaccaag aganaancgt ctgccaacac caaccgcccc agcctggcg ggcacganag 240
gaaactgggt accaatctgc agaattctna gaggaanaag cnagggggccc gcgcgtnaga 300
cagagctgga tatgangcca gaccatggac nctacncccn ncaatncana cgggaactgcg 360
gaagatggan gaccncgcac nngatcaggc cngctnncca nccccccacc cctatgaatt 420
attcccgctg aangaatctc tgannggctt ccannaaagc gcctcccncn cnaacgnaan 480
tncaacatng ggattanang ctgggaactg naaggggcaa anctnnnaat atccccagaa 540
acaanctctc ccnaanaaac tggggcncct catngtggg accaactatt aactaaaccg 600
cacgcgaag aantataaaa ggggggcccc tcncggmng acccctcttt gtcccttaat 660
ganggttatc cnccttgctg accatggtnc cennctctgt ntgnatgttt cncctcccc 720
cncctatnt cnagccgaac tcnnatttnc ccgggggtgc natcnantng tncnctttt 780
ttngttgncc cngcccttcc cngcggaacn cgtttcccg tntantaaagg caccggggm 840
aagggtgntt ggcgccctcc ctccc 865

```

&lt;210&gt; 72

&lt;211&gt; 560

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

```

<222> 83, 173, 183, 186, 209, 211, 215, 255, 321, 322, 323, 335,
344, 357, 361, 368, 394, 412, 415, 442, 455, 469, 472, 475,

```

487, 513, 522, 528, 531, 534, 546

<223> n = A,T,C or G

<400> 72

```
cctggacttg tcttggttcc agaacctgac gacccggcga cggcgacgtc tcttttgact 60
aaaagacagt gtccagtgtc ccngcctagg agtctacggg gaccgcctcc cgcgcgcgca 120
ccatgcccaa cttctctggc aactggaaaa tcatccgato ggaaaacttc gangaattgc 180
tcnaantgct ggggggtgaat gtgatgctna ngaaanattgc tgtggctgca gcgtccaagc 240
cagcagtgga gatcnaacag gaggggagaca ctttctacat caaaacctcc accacogtgc 300
gcaccacaaa gattaacttc nnnngttggg aggannttga ggancaaact gtggatnnga 360
ngcctgtnaa aacctgtgtga aatggggagaa tganaataaa atggtgtgtg ancanaaact 420
cctgaaagga gaagcctccc anaactcctg gaccngaaaa actgaccncn cnatngggga 480
actgatnctt gaacctgaa cgggcgggat ganocctttt tnttgccncc naanggggtc 540
tttccntttc cccaaaaaaa                    560
```

<210> 73

<211> 379

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 8, 17, 18, 21, 26, 29, 30, 32, 53, 56, 67, 71, 81, 102, 104, 111, 112, 114, 119, 122, 124, 125, 134, 144, 146, 189, 190, 214, 215, 219, 220, 235, 237, 246, 280, 288, 302, 310, 313, 319, 322, 343, 353, 354

<223> n = A,T,C or G

<400> 73

```
ctggggancc ggcggtngcc nccatntcnn gncgcgaagg tggcaataaa aanccnctga 60
aaccgcncaa naaacatgcc naagatatgg acgaggaaga tngngcttcc nngnacaanc 120
gnanngagga acanaacaaa ctcnangagc tctcaageta atgcgcgggg gaagggggccc 180
ttgggccacnn gtggaattaa gaaatctggc aaanngtann tgttcctgtg gcctnangag 240
ataaangacc ctttatttca tctgtattta aacctctctn ttccctgnca taactctctt 300
tnccacgtan agntggaant antgtgttgo ttggactgtt gtncatttta gannaacact 360
ttgttcaaaa aaaaaataa                    379
```

<210> 74

<211> 437

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 145, 355

<223> n = A,T,C or G

<400> 74

```
actagttcag actgccacgc caaccccaga aaatacccca catgccagaa aagtgaagtc 60
ctagggtgttt ccatctatgt ttcaatctgt ccatctacca ggccctogcg taaaaacaaa 120
acaaaaaaac gctgcccaggt tttnaagaca gttctgtgct caaaaccatc aggatctctgc 180
caccagggtt cttttgaaat agtaccacat gttaaaggga atttgggttt cacttoactc 240
aatcaactgaa ttgtcaggct ttgattgata attgtagaaa taagtagcct tctgttgtgg 300
gaataagtta taatcaagtat tcatctcttt gttttttgtc actcttttct ctctnattgt 360
gtcatattgta ctgtttgaaa aatatctctt ctataaaatt aaactaacct gccttaaaaa 420
aaaaaaaaaa aaaaaaa                    437
```

<210> 75

<211> 579  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 440, 513, 539, 551  
 <223> n = A,T,C or G

<400> 75  
 ctccgtcgcc gccaaatga tgtcgggggc gccctccgcc acgcagccgg ccaccgcga 60  
 gaccocagac atcgccgacc aggtgaggtc ccagcttgaa gagaaagaaa acaagaagtt 120  
 ccctgtgttt aaggccgtgt cattcaagag ccaggtggtc gcggggacaa actaactcat 180  
 caaggtgcac gtcggcgacg aggaacttcgt acacctgoga gtgttccaat ctctccctca 240  
 tgaaaaacaag cccttgacct tatctaacta ccagaccaac aaagccaagc atgatgagct 300  
 gacctatttc tgatcctgac ttgggacaag gcccttcagc cagaagactg acaagtcatt 360  
 cctccgtcta ccagagcgtg cacttgtgat cctaaaataa gcttcattct cgggctgtgc 420  
 ccttgggggtg gaagggggcan gatctgcact gcttttgcat ttctctctct aaatttcatt 480  
 gtgttgattc ttctcttcca ataggtgact ttnattactt tcagaatatt ttccaaatna 540  
 gatattttt naaaatcctt aaaaaaaaaa aaaaaaaaaa 579

<210> 76  
 <211> 666  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 411, 470, 476, 491, 506, 527, 560, 570, 632, 636, 643, 650,  
 654, 658  
 <223> n = A,T,C or G

<400> 76  
 gtttatccta tctctccaac cagattgtca gctccttgag ggcaagagcc acagtatatt 60  
 tccctgtttc ttccacagtg cctaataata ctgtggaact aggtttttaa aattttttta 120  
 ttgatgtttt tatgggcagg atggcaacca gaccattgtc tcagagcagg tgctggctct 180  
 ttctgggcta ctccatgttg gctagcctct ggtaacctct tacttattat ottcaggaca 240  
 ctactacag ggaccaggga tgatgcaaca tctctgtctt tttatgacag gatgtttgct 300  
 cagcttctcc aacaataaaa agcacgtggt aaaacacttg cggatattct ggactgtttt 360  
 taaaaaatat acagttttaco gaaaatcata ttatcttaca atgaaaagga ntttatagat 420  
 cagccagtgta acaacctttt cccaccatcac aaaaattcct ttcccccgaan gaaaanggct 480  
 ttctcaataa ncttcaacttt ottaanatct tacaagatag ccccganac ttatcgaac 540  
 tcatttttagg caaatatgan ttttattgtt cgttacttgt ttcaaaattt ggtatttgta 600  
 atatcaatta ccaccccatc ctcccatgaa anaanggga aanggtgaan ttctaanagc 660  
 ottaaa 666

<210> 77  
 <211> 396  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 31, 54, 125, 128, 136, 163, 168, 198  
 <223> n = A,T,C or G

<400> 77  
 ctgcagcccg ggggatccac taatctacca nggttatttg gcagctaatt ctanatttg 60

```

atcattgccc aaagtgcac ttgctggtct cttgggattt gcccttgaa aggtatcata 120
catanganta tgccanaata aattccattt ttttgaaat canctcctntg gggctggttt 180
tggtccacag cataacangc actgcctcct tacctgtgag gaatgcaaaa taaagcatgg 240
attaagttag aaggggagact ctcagccttc agcttccctaa attctgtgtc tgtgactttc 300
gaagtttttt aaacctctga attttgacac atttaaaatt tcaagtgtag ttttaaaataa 360
aatacttcta atgggaacaa aaaaaaaaaa aaaaaa

```

```

<210> 78
<211> 793
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 309, 492, 563, 657, 660, 703, 708, 710, 711, 732, 740, 748,
758, 762, 765, 787
<223> n = A,T,C or G

```

```

<400> 78
gcattcctagc cgccgactca cacaaggcag gtgggtgagg aaatccagag ttgccatgga 60
gaaaattcca gtgtcagcat tcttgtctct tgtggccctc tctacactc tggccagaga 120
taccacagtc aaacctggag ccaaaaagga cacaaggac tctgaccca aactgcccca 180
gacctctctc agagggttgg gtgaccaact catctggact cagacatag aagaagctct 240
atataaatcc aagacaagca acaaaccttt gatgattatt catcaacttg atgagtgccc 300
acacagtcna gctttaaaga aagtgtttgc tgaaaataaa gaaatccaga aattggcaga 360
gcagtttgct ctcctcaatc tggtttatga aacaactgac aaacaccttt ctctgatgg 420
ccagtatgtc ccaggattat gtttgttgac ccatctctga cagttgaagc cgatatcctg 480
ggaagatatt cnaaccgtct ctatgcttac aaactgcaga tacgctctgt tgcttgacac 540
atgaaaaaagc tctcaagttg ctnaaaatga attgtaagaa aaaaaatctc cagccttctg 600
tctgtcgggt tgaaaattga aaccagaaaa atgtgaaaaa tggctatttg ggaacanatn 660
gacacctgat taggttttgg ttatgttcac cactattttt aanaaaanan nttttaaaat 720
ttggttcaat tntcttttn aaacaatntg tttctacntt gnganctgat ttctaaaaaa 780
aataatnttt ggc

```

```

<210> 79
<211> 456
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 89, 195, 255, 263, 266, 286, 353, 384, 423, 425, 436, 441
<223> n = A,T,C or G

```

```

<400> 79
actagtatgg ggtgggaggc ccacaccttc tcccttaggc gctgttcttg ctocaaaggg 60
ctccgtggag agggactggc agagctgang ctgacctggg ctggggatcc cactcttctt 120
gcagctgttg agcgcaacta accaactggc atgccccac cctgtctctc cgcacccgct 180
tccctccgac ccangacca ggctacttct cccctcctct tgccctccctc gtgccccgct 240
tgccctctgat cgtangaatt gangantgtc ccgacctgtg gctganaatg gcacgtggca 300
ggggctggaa atgggtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gncccccccc 360
tgcaagaccg agattgagg aaancatgtc tgctgggtgt gaccatgttt cctctccata 420
aantnccccct gtgacnctca naaaaaaaaa aaaaaa

```

```

<210> 80
<211> 284
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 283
<223> n = A,T,C or G

<400> 80
ctttgtacct ctagaaaaa tagtgattgt gtcatgaaac ttgagtttaa attttatata 60
taaaactaaa agtaatgtct accttagcaa cacatactaa aattggaacc atactgagaa 120
gaatagcatg acctccgtgc aaacaggaca agcaaatattg tgatgtgttg attaaaaaga 180
aataaaataa tgtgtatatg tgtaacttgt atgtttatgt ggaatacaga ttgggaaata 240
aatgtatatt cttactgtga aaaaaaaaaa aaaaaaaaaa aana 284

<210> 81
<211> 671
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 388, 505, 600, 603, 615, 642, 644, 660
<223> n = A,T,C or G

<400> 81
gccaccaaca ttccaagcta ccctgggtac ctttgtgcag tagaagctag tgagcatgtg 60
agcaagcggt gtgcacacgg agactcatcg ttataattta ctatctgcc aagtagagaa 120
gaaaggctgg ggatatttgg gtgggotttg ttttgatttt ttgcttggtt gtttgttttg 180
tactaaaaaa gtattatctt ttgaatatcg tagggacata agtatataca tgttatccaa 240
tcaagatggc tagaatgggt cctttctgag tgtctaaaac ttgacacccc tggtaaatct 300
ttcaacacac ttccactgcc tgcgtaatga agttttgatt catttttaac cactggaatt 360
tttcaatgcc gtcattttca gttagatnat ttgacacttt gagattaaaa tgccatgtct 420
atttgattag tcttattttt ttatttttac aggcattatca gtctcaactgt tgggtgtcat 480
tgtgacaaag tcaataaaac ccccnaggac aacacacagt atgggatcac atattgtttg 540
acattaaagt ttggccaaaa aatggtgcat gtgttttaac tcgacttgct aaatcaatan 600
canaaaggct ggctnataat gttggtggtg aaataattaa ntantaacca aaaaaaaaaa 660
aaaaaaaaaa a 671

<210> 82
<211> 217
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 35
<223> n = A,T,C or G

<400> 82
ctgcagatgt ttcttgaatg ctttgtcaaa ttaanaaagt taaagtgcaa taatgtttga 60
agacaataag tgggtgtgta tcttgtttct aataagataa acttttttgt ctttgcttta 120
tcttattagg gagttgatg tcagtgtata aaacatactg ttggtgataa caggcttaat 180
aaattcttta aaaggaaaaa aaaaaaaaaa aaaaaaa 217

<210> 83
<211> 460
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 104, 118, 172, 401, 422, 423, 444, 449
<223> n = A,T,C or G

<400> 83
cgcgagtggg agcaccagga tctcgggctc ggaacgagac tgcacggatt gttttaagaa 60
aatggcgagc aaaccagaca tgggggaaat cgccagcttc gatnaggcca agctgaanaa 120
aacggagacg caggagaaga acaccctgcc gaccaagag accattgagc angagaagcg 180
gagtgaaatt tcttaagatc ctggaggatt tcttaccocc gtctctcttc agaccccgagt 240
cgtgatgtgg aggaagagcc acctgcaaga tggacacgag ccacaagctg cactgtgaac 300
ctgggacactc cgccgcgatg ccaccggcct gtgggtctct gaagggagccc ccccaaatcg 360
gactgccaaa ttctccgggt tgcgccggga tattatacaa nattatttgt atgaataatg 420
annataaaac acacctcggt gcancaaaaa aaaaaaaaaa 460

<210> 84
<211> 323
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 70, 138, 178, 197, 228, 242, 244, 287, 311
<223> n = A,T,C or G

<400> 84
tggtggatct tggctctgtg gagctgctgg gacgggatct aaaagactat tctggaagct 60
gtgggtccaan gcattttggc ggcttaacgg gtcccggaac aaaggacacc agctctctaa 120
aattgaagtt tacccganat aacaatcttt tgggcagaga tgctattttt aacaaacncc 180
gtccctgcgc aacaaacnaa aatctctggg aaataccggc catgaacntg ctgtctcaat 240
cnanacatctc tctagctgac cgatcatatc gtcccagatt actacanatc ataataattg 300
atttctctga naaaaaaaaaa aaa 323

<210> 85
<211> 771
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 63, 426, 471, 497, 521, 554, 583, 586, 606, 609, 615, 652,
686, 691, 694, 695, 706, 713, 730, 732, 743, 751
<223> n = A,T,C or G

<400> 85
aaactgggta ctcaacactg agcagatctg ttctttgagc taaaaacat gtgctgtacc 60
aanagtttgc toctggctgc tttgatgtca gtgctgtac tccacotctg cggcgaaatca 120
gaagcaagca actttgactg ctgtcttgga tacacagacc gtattcttca tctcaaatatt 180
atttggggct tcacacggca gctggccaat gaaggctgtg acatcaatgc tatcatcttt 240
cacacaaga aaaagtgtgc tgtgtgcgca atccaaaaac agacttgggt gaatatatt 300
gtgcgtctcc tcagtaaaaa agtcaagaac atgtaaaaac tgtggctttt ctggaatgga 360
attggacata gcccaagaac agaaagaact tgcgtgggtt ggaggtttca ctgtcacatc 420
atgganggtt tagtgcttat cttatttgtg cctcctggac ttgtccaatt natgaagtta 480
accatatctg atcatanttt gctttgttta acatcacatt naaattaaac tgtattttat 540
gttatttata gctntaggtt tctgtgtttt aattttttat acnaantttc ctaaaactatt 600
ttggtntant gcaanttaaa aattatattt ggggggggaa taaatatttg antttctgca 660
gccacaagct ttttttataa aaccantaca nccnngttaa atgtgntggc cnaatgggtt 720
tttgcttttn antagaaaaa ttnttgaac natttgaaaa aaaaaaaaaa a 771

```

<210> 86  
 <211> 628  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 162, 249, 266, 348, 407, 427, 488, 518, 545, 566, 569, 597,  
 598, 611, 617, 621, 624  
 <223> n = A,T,C or G

<400> 86  
 actagttgttc tttacatttt tgaagaagtat tatttttgtc caagtgtctta tcaactaaac 60  
 cttgtgtgttag gtaagaatgg aatttatttaa gtgaatcagt gtgaccttc ttgtcataag 120  
 attatctttaa agctgaagcc aaaatatgct tcaaaagaaa angactttat tgttcattgt 180  
 agttcataca ttcaaaagcat ctgaactgta gtttctatag caagccaatt acatccataa 240  
 gtggagaang aaatagatta atgtcnaagt atgattgggt gagggagcaa ggttgaagat 300  
 aatctgggggt tgaatttttc tagttttcat totgtacatt tttagttna catcagattt 360  
 gaaatattaa tgtttacctt tcaatgtgtg gtatcagctg gactcantaa ccccccttc 420  
 ttccctnngg gatggggaat ggattattgg aaaatggaaa gaaaaaagta cttaagcct 480  
 tccctttnca gtttctggct cctaccctac tgatttanc agaataagaa aacattttat 540  
 catcntctgc tttattccca ttaatnaant tttgatgaat aaatctgctt ttatgcnnac 600  
 ccaaggaatt nagtgnnttc ntcnttgg 628

<210> 87  
 <211> 518  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 384, 421, 486  
 <223> n = A,T,C or G

<400> 87  
 ttttttattt tttttagaga gtagttcago tttttttat aaatttattg cctgttttat 60  
 tataacaaca ttatactggt tatgttttaa tacatatggt tcaaatgtga taatacatca 120  
 agtagtacag ttttaaaatt ttatgcttaa aacaagtttt gtgtaaaaaa tgcagatata 180  
 ttttcatggt caaatcaatt ttttaagtcac cctaaaaaatt gatttttttt tgaattttaa 240  
 aaacacattt aatttcaatt totctcttat ataaccttta ttactatagc atgtgtttcca 300  
 ctacagttta acaatgcagc aaaattccca tttcacggta aattgggttt taagcgccaa 360  
 ggttaaaatg ctttgaggat cctnaatacc ctttgaactt caaatgaagg ttatggttgt 420  
 naatttaacc ctcatgccat aagcagaagc acaagtttag ctgcattttg ctctaaacctg 480  
 taaaanogag ccccccgttg aaaaagcaaa agggacc 518

<210> 88  
 <211> 1844  
 <212> DNA  
 <213> Homo sapiens

<400> 88  
 gagacagtga atcctagtag caaaggattt ttggcctcag aaaaagtgtg tgattatttt 60  
 tatttttatt tattttttcg gactccgtct caaaaaaaaa aaaaaaaaaga gaattcacia 120  
 ggtatttgcg aaagcatttt gagctgcttg gaaaaaggga agtagttgca gttaggtttc 180  
 ttccatcttc ttggtgctgg gaagccatat atgtgtcttt tactcaagct aaggggtata 240  
 agcttatgtg ttgaatttgc tactatcata ttccacatat tctcacaata agagaatttt 300  
 gaaatagaaa tatcatagaa catttaagaa agtttagtat aaataatatt ttgtgtgttt 360

```

taatcccttt gaagggatct atccaaagaa aatattttac actgagctcc ttctacacg 420
tctcagtaac agatcctgtg ttagtctttg aaaatagctc atttttttaa tgtcagtgag 480
tagatgtagc atacatatga tgtataatga cgtgtattat gtaacaatg tctgcagatt 540
ttgtaggaat acaaaacatg gcccttttta taagcaaaac gggccaatga ctagaataac 600
acatagggca atctgtgaaat atgtattata agcagcattc cagaaaaagta gttgggtgaa 660
taattttcaa gtcaaaaagg gatattgaaa gggaaattatg agtaacctct atttttttaag 720
ccttgctttt aaatttaaag ctacagccat ttaagccttg aggataataa agccttgagag 780
taataatgtt aggttagcaa aggttttagat gtatcacttc atgcattgcta ccattgatgt 840
aatgcagctc ttcgagtcatt ttctgggtcat tcaagatatt cacccttttg cccatagaaa 900
gcacccatcc tcacctgctt actgacattg tcttagctga tcacaagatc attatcagcc 960
tcacattatc cttaactgtat ataaaaatca gagttttata ttttcccttc ttogtttttc 1020
accattatca aaacctaaat ttgttttttg agatggaatg caaagtaatc aagtgttcgt 1080
gctttccact agaagggtgt ggtcctgaag gaaagaggtc cctaaatatt ccccaacctg 1140
ggtgctcctc cttccctggt accctgacta ccagaagtc ggtgctagag cagctggaga 1200
cgtgcagcag cctgtgcttc cacagatggg ggtgctgctg caacaaggct ttcaatgtgc 1260
ccatccttag gggagaagct agatcctgtg cagcagcctg gtaagtctgt agggagttcc 1320
attgctcttc ctgctgctgt cctttgcttc tcaacggggc tcgctctaca gctagagca 1380
catgcagctc actgtgacct ctgcttatgc atgagggtta aattaacaac cataacctc 1440
atttgaagtt caaaggtgta ttcaggatcc tcaaacattt ttaaccttgc cgtcttaaac 1500
ccaatttacc ggaattggg aattttgctg cattgtttaa ctgtagtga accatgcta 1560
tagtaataaa ggttatataa gagagaat gaaattaaat gtgtttttaa atttcaaaaa 1620
aaaatcaatc tttaggatga cttaaaaaatt gatttgccat gtaaatgta tctgcatttt 1680
ttacacaaaa cttgttttaa gcataaaaatt ttaaaactgt actactgtat gtattatata 1740
ttttgaacca tatgtatttaa accataaaca gtataatgtt gttataataa aacaggcaat 1800
aaatttataa ataaaagctg aaaaaaaaaa aaaaaaaaaa aaaa 1844

```

&lt;210&gt; 89

&lt;211&gt; 523

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 288, 352, 369, 398, 475, 511, 513

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 89

```

tttttttttt tttttttagt caatccacat ttattgatca cttattatgt accaggcact 60
gggataaaga tgactgttag tcaactcacag taaggagaag aactagcaaa taagacgatt 120
acaattatgat gtagaaaatg ctaagccagag gatatagaaa ggtcctattg ggtcctcttg 180
tcacctgtgc tttccacatc cctacccttc acaggccttc cctccagctt cctgcccccg 240
ctccccactg cagatccctt gggattttgc cttagagctaa acgagganat gggccccctg 300
gccctggcat gacttgaacc caaccacaga ctgggaaagg gaggcctttc anagtggatc 360
actttgatna gaaaacacat aggggaattg agagaaantc cccaaatggc caccogtgct 420
ggtgctcaag aaaagtttgc agaattggata aatgaaggat caagggaatt aatanatgaa 480
taattgaatg gtggtcatt aagaatgact ncnttgaatg acc 523

```

&lt;210&gt; 90

&lt;211&gt; 604

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 563

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 90



```

ccagtgtggt ggaatgcaaa gattaccccg gaagctttcg agaagctggg attccctgca 60
gcaaaggaaa tagccaatat gtgtcggttc tatgaaatga agccagacgc agatgtcaat 120
ctcaccacacc aactaaatcc caaagtcaaa agcttcagcc agtttatctc agagaaccag 180
gggagccttc aagggcatgt agaaaatcag ctgttcagat aggcctctcg accacacagc 240
ctctttctct totgatctct ttctctctta cggcacaaca ttcatgtttg acagaacatg 300
ctggaatgca attgttttga acaccgaagg atttctcgcg gtcgcctctt cagtaggaag 360
cactgcattg gtgataggac acggtaattt gattcacatt taacttgcta gttagtata 420
aggggtggta cacctgtttg gtaaaatgag aagcctcgga aacttgggag cttctctctc 480
accactaatg gggagggcag attattactg ggattttctc tggggtggaat taatttcaag 540
cctaatttgc tgaatttccc ctnggcagcg tccagtttcc tcaactgcat tgcaaaattc 604
cccc

```

&lt;210&gt; 91

&lt;211&gt; 858

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

```

<222> 570, 591, 655, 664, 667, 683, 711, 759, 760, 765, 777, 787,
792, 794, 801, 804, 809, 817, 820

```

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 91

```

tttttttttt ttttttttta tgattattat tttttttatt gatctttaca tcctcagttg 60
tggcagagtt totgatgctt aataaacatt tgtttctgat agataagtgg aaaaatttgt 120
catttcctta ttcaagccat gcttttctgt gatattctga tcctagtgtg acatacagaa 180
ataaatgtct aaaacagcac ctcgattctc gtctataaca ggactaaagt cactgtgatc 240
ttaaataaag ttggctaaaa tgggacatga gtggagtgag tcacacttca gcgaagaag 300
agaatctctc gtataatctc accaggagat tcaacgaatt ccaccaactt ggaactagtgg 360
atcccccgcg ctgcagggaat tcgatataca gcttatcgat accgtogacc tcgagggggg 420
gcccggtacc caattgcgcc tatagtgaat gtattacgc gcgctcactg gccgtcgttt 480
tacaacgtcg tgactgggaa aaccctggcg ttatcccaat taatgcctt gcagcacatc 540
cccccttcgc cagctggcgt aatagcgaan agcccgacc gatcgccctt ncaacagttg 600
cgcagcctga atggcgaatg ggacgcgcgc tgtagcggcg cattaaagcg cggcnggggtg 660
tggngngtcc cccacgtgac cgntacaact ggcagcgccct taocgcggtc ntctcgctttc 720
ttcctctctc ttctcgacc gttcgcggg ttctcccgnn agctnttaat cgggggntc 780
cctttanggg tncnaattaa nggnttacng gacctnngan cccaaaaact ttgattaggg 840
ggaaggtccc cgaagggg

```

858

&lt;210&gt; 92

&lt;211&gt; 585

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

```

<222> 317, 319, 320, 321, 325, 327, 328, 330, 331, 332, 460, 462,
483, 485, 487, 523, 538, 566, 584

```

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 92

```

tttgaatctc ctggtgagat tatacaggag attctotttc ttogctgaag tgtgactacc 60
tccactcatg tccattttta gccaaagcta tttaagatca cagtgaaact agtctgttta 120
tagacagaaa tcgaggtgct gtttttagaca tttatttctg tatgttcaac taggatcaga 180
atatcacaga aaagcatggc ttgaataaag aaatgacaat tttttccaat tatctgatac 240
gaacaaatgt ttattaaaga tcagaaactt tgccaaact gaggatgtaa agatcaataa 300
aaaaataat aatcatnann naaanannan nngaaggcgc gccgccaccg cgggtggagct 360

```

```
ccagctttttg ttcccttttag tgagggttaa ttgcgcgctt ggcgttaatc atggtcatag 420
ctgttttctgt tgtgaaattg ttatccggct cacaattccn cnaacatac gagccgggaa 480
gcntnangtg taaaagcctg ggggtgccta attgagttag ctactacaca ttaattngnt 540
tgcgctccac ttgcgcgctt ttccantccg ggaaacctgt tcgnc 585
```

```
<210> 93
<211> 567
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<222> 82, 158, 230, 232, 253, 266, 267, 268, 269, 270, 271, 272,
273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284,
285, 286, 287, 295, 303, 307, 314, 349, 352, 354, 356, 366,
369, 379, 382, 386, 393, 404, 427, 428, 446, 450, 452
<223> n = A,T,C or G
```

```
<221> misc_feature
<222> 453, 454, 459, 462, 480, 481, 483, 488, 493, 501, 509, 511,
512, 518, 520, 525, 526, 532, 541, 557
<223> n = A,T,C or G
```

```
<400> 93
cggcagtggt gctgtctgcg tgtccacctt ggaatctggc tgaactggct gggaggacca 60
agactgcggc tgggggtggc anggaaggga accgggggct gctgtgaagg atcttggaa 120
ttccctgtac ccaacctccc ottgttccat gtttgtanag gaaccttctg ccggccaagc 180
coagtttctc tgtgtgatac actaatgtat ttgctttttt tgggaaatan anaaaaatca 240
atataattgc tantgtttct ttgaannnnn nnnnnnnnnn nnnnnnnngg gggngcggcc 300
cncggngga aacnccccc ttgttccct ttaattgaaa ggttaattng cncnctggc 360
gttaancntc gggccaaanc tngttncccg tngtgaattt gttnatcccc tcccaaatcc 420
cccccnccc ttccaaaacc ggaaancccn anntgttna anccggggg gttgctaan 480
ngnaattnaa ccnaaccccc nttaaatng nntttgcn cnacnngccc cnccttccca 540
nttcggggaa aacctntcc gtgcca 567
```

```
<210> 94
<211> 620
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<222> 169, 171, 222, 472, 528, 559, 599
<223> n = A,T,C or G
```

```
<400> 94
actagtcaaa aatgctaaaa taatttgga gaaaaatttt ttaagttagt gttatagttt 60
catgtttatc ttttattatg ttttgtgaag ttgtgtcttt tcaactaatta cctatactat 120
gccaatattt ccttatatct atccataaca ttataactac atttgtaana naatatgcac 180
gtgaaacctt acaatttata aggtaaaaat gaggtttcca anatttaata atctgatcaa 240
gtttctgtta ttccaaaata gaatggactt ggtctgttaa gggctaaggga gaaggaggaa 300
ataagggttaa aagttgttaa tgacccaaaa ttctaaaaaga aatgcacaaaa aaaagtttat 360
tttcaagcct togaactatt taaggaaagc aaaatcattt cctaaatgca tatcatttgt 420
gagaatttct cattataatc ctgaatcatt catttcaata aggcctatgt tnaactcgat 480
atgtctctaa gaaagtaact tttcatggtc caaacctggt tgccatantt gggtaaaagg 540
tttccttaa gtgtgaaant atttaaatg aaattttcct ctttttaaaa attctttana 600
aggggttaag gtgttggga 620
```

```
<210> 95
<211> 470
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 61, 67, 79, 89, 106, 213, 271, 281, 330, 354, 387, 432, 448
<223> n = A,T,C or G

<400> 95
ctogaccttc tctgcacagc ggatgaaccc tgagcagctg aagaccagaa aagccactat 60
nactttntgc ttaattcang agcttacang atttctcaaa gagtgngtoc agcatccttt 120
gaaacatgag ttcttaccag cagaagcaga cctttacccc accacctcag ctccaacagc 180
agcagggtgaa acaacccatc cagcctccac cttaggaaat atttgttccc acaaccaagg 240
agccatgccca ctcaaaaggtt caacaacotg naaacacaaa nattccagag ccaggctgta 300
ccaaggtccc tgagccaggg ctgtaccaan gtcctgagc caggttgtag caangtccct 360
gagccaggat gtaccaaggt ccotgancca ggttgccaa ggtccctgag ccaggctaca 420
ccaagggcct gngccaggca gcatcaangt ccctgaccaa ggttatcaa 470

<210> 96
<211> 660
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 299, 311, 360, 426, 538, 540, 542, 553, 563, 565, 592, 603,
604, 618, 633, 647, 649, 651, 653
<223> n = A,T,C or G

<400> 96
tttttttttt tttttttttt ggaattaaaa gcaatttaat gagggcagag caggaaacat 60
gcatttcttt tcatttcgaat ctccagatga accctgagca gccgaagacc agaaaagcca 120
tgaagacttt ctgcttaatt caggggctta caggattctt cagagtgtgt gtgaacaaaa 180
gctttatagt acgtattttt aggatacaaa taagagagag actatggctt ggggtgagaa 240
tgtactgatt acaaggtcta cagaacaatta agacacagaa acagatggga agagggtgnc 300
cagcatctgg nggttggtct ctcaagggtc tgtctgtgca ccaaattact totgcttggc 360
cttctgtgtg gctgggcttg gagtgaacct tgaaggacat ggctctggta cotttgtgta 420
gcctgnacaa ggaacttttg tgtatccttg ctcaggaaat ttgatggcac ctggctcagg 480
aaacttgatg aagccttggt caagggacct tgatgtctgc tggctcaggg acotttgngn 540
ancctgggct canggaacct tgnncnaacc ttggcttcaa gggacccctg gnacatcctg 600
gcnnagggac ccttggnncc aaccctgggc ttnagggacc ctttggttnc nanccttggc 660

<210> 97
<211> 441
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 12, 308
<223> n = A,T,C or G

<400> 97
gggaccatac anagtattcc tctcttcaca ccaggaccag cactgtgtgc agcatgagtt 60
cccagcagca gaagcagccc tgcattccac cccctcagct tcagcagcag caggtgaac 120
```

```

agccttgcca gctccacct caggaacct gcattcccaa aaccaaggag ccctgccacc 180
ccaagggtgcc tgagccctgc caccocaaag tgcctgagcc ctgccagccc aaggttccag 240
agccatgcca ccccaagggt cctgagccct gcccttcaat agtccactcca gccaccagccc 300
agcagaanac caagcagaag taatgtggto cacagccatg cccttgagga gccggccacc 360
agatgctgaa tcccctatcc cattctgtgt atgagtccea ttgtccttgc aattagcatt 420
ctgtctcccc caaaaaaaaa a

```

```

<210> 98
<211> 600
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 295, 349, 489, 496, 583
<223> n = A,T,C or G

```

```

<400> 98
gtattctctct cttcacacca ggaccagcca ctgttgagc atgagttccc agcagcagaa 60
gcagccctgc atcccacccc ctgagcttca caaggagccc tgccacccca aggtgcctga 120
tccacctcag gaacctatgca tcccocaaac ccaaggagccc tgccacccca aggtgcctga 180
gccctgccac cccaaagtgc ctgagccctg ccagcccaag gttccagagc catgccaccc 240
caagggtgct gagccctgcc ctccaatagt cactccagca ccagcccagc agaanaacca 300
cagaagtaaa tgtggtccac agccatgccc ttgaggagcc ggccaccana tgctgaatcc 360
cctatcccat tctgtgtatg agtccatttt gcccttgaat tagcattctgt tctcccccac 420
aaaagaatgt gctatgaagc tttctttcct acacactctg agtctctgaa tgaagctgaa 480
ggtcttaant acagantctag ttttcagctg ctcagaattc tctgaagaaa agatttaaga 540
tgaaggacaa atgattcagc tccttattac cccattaaat tcnctttcaa ttocaaaaaa 600

```

```

<210> 99
<211> 667
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 345, 562, 635
<223> n = A,T,C or G

```

```

<400> 99
actagtgaact gagttcctgg caaagaaatt tgacctggac cagttgtataa ctcattgttt 60
accatttaaa aaaatcagtg aaggatttga gctgctcaat tcaggacaaa gcattogaac 120
ggtcctgaag ttttgagatc caaagtggca ggaggctctgt gttgtcatgg tgaactggag 180
tttctctgtg gagagttccc tcactctgaaa tcatgtatct gtctcacaaa tacaagcata 240
agtagaagat ttgttgaaga catagaaccc ttataaagaa ttattaaact ttataaacat 300
ttaaagcttt gtgagcaact gggaattagt ataatacaaa tggttnatatt ttgtatttac 360
atatttgaag gctataattg tatcttttaa gaaaacatac cttggatttc tatgttgaaa 420
tggagatttt taagagtttt aaccagctgc tgcagatata ttactcaaaa cagatatagc 480
gtataaagat atagtaaatg catctcctag agtaatatct acttaacaca ttggaaacta 540
ttatttttta gatttgaata tnaatgttat tttttaacaa cttgttatga gttacttggg 600
attacatttt gaaatcagtt cattccatga tgcanaattac tgggattaga ttaagaagaa 660
cggaaaa

```

```

<210> 100
<211> 583
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 404, 506, 514, 527, 528, 538, 548, 556, 568, 569
<223> n = A,T,C or G

<400> 100
gttttgtttg taagatgac acagtcacgt tacactgacg taaaggacat atatataacc 60
ctttaaaaaa aaaatcactg cctcattctt atttcaagat gaatttctat acagactaga 120
tgttttcttg aagatcaatt agacattttg aaaatgattt aaagtgtttt ccttaagtgt 180
ctctgaaaac aagtttcttt tgtagtttta accaaaaaag tgcctttttt gtoactggat 240
tctcctagca ttcattgattt ttttttoata caatgaaatt aaaattgcta aaatcatgga 300
ctggctttct gggttgattt caggtaagat gtgtttaagg ccagagcttt tctcagttat 360
tgattttttt ccccaattatt tgatttttta aaaatataca catnggtgct gcatttttat 420
ctgtcgtggt aaaattctgt catatttcac tctagcctt ttagtattgg caaatcatat 480
tttactttta cttaaagcat ttggttattt ggantatctg gtcttannct aaaaaaanta 540
attctatnaa ttgaantttt ggtactcnnc catatttgga tcc 583

<210> 101
<211> 592
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 218, 497, 502, 533, 544, 546, 548, 550, 555
<223> n = A,T,C or G

<400> 101
gtggagacgt acaaaagagca gccgctcaag acacctggga agaaaaagaa aggcaagccc 60
gggaaacgca agggagcagga aaagaaaaaa cggcgaaact gctctgctgt gttagactct 120
ggagtgactg ggagtggggt agaagggggac cacctgtctg acacctccac aacgtgctgt 180
gagctcgatt cacggaggca ttgaaatttt cagcaganac cttccaagga catattgcag 240
gattctgtaa tagtgaacat atggaaaagta ttgaaaatat ttattgtctg taaatactgt 300
aaatgcattg gaataaaaact gtctcccca ttgctctatg aaactgcaca ttggtcattg 360
tgaatatatt tttttttgcc aaggctaatc caattattat tatcacattt accataattt 420
attttgtcca ttgatgtatt tattttgtaa atgtatcttg gtgctgtgta atttctatat 480
tttttgtaca taatgcnttt anatacacct atcaagtttg ttgataaatg acncaatgaa 540
gtgncnncan ttgngggttg aatttaatga atgcctaatt ttattatccc aa 592

<210> 102
<211> 587
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 91, 131, 256, 263, 332, 392, 400, 403, 461, 496, 497, 499,
510, 511, 518, 519, 539, 554, 560, 576
<223> n = A,T,C or G

<400> 102
gctcctaagc acttagacta catcagggaa gaacacagac cacatccctg tctctcatg 60
gcttatgttt tctggaagaa agtggagacc nagtcccttg ctttagggct ccccgctgg 120
gggctgtgca ntcoggtcag ggcggaagg gaaatgcacc gctgcatgtg aacttacagc 180
coaggcggtat gcccttccc tttagactac ctggcctoct gcatcccttc gctcattgtt 240
cctccacctt caaaaatag aanaacccca tgggcaccag cctctgccct ggggaaccaa 300
ggcagccttc caaaactcag gggctgaagc anactattag ggcaggggct gactttgggt 360

```

```

gacactgccc attccctctc agggcagctc angtcacccn ggnctcttga acccagcctg 420
ttcctttgaa aaagggcaaa actgaaaagg gcttttctta naaaaagaaa aaccagggaa 480
ctttgcccagg gcttcnntnt taccaaaaacn ncttctcnng gattttttaa tccccattng 540
gcctccactt accnggggcn atgccccaaa attaanaatt tcccatc 587

```

```

<210> 103
<211> 496
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc feature
<222> 2, 17, 66, 74, 82, 119, 164, 166, 172, 200, 203, 228, 232,
271, 273, 415, 423, 445, 446, 473
<223> n = A,T,C or G

```

```

<400> 103
anaggactgg cccactnctg tctctctcgt cctacctatc aatgcccaac atggcagaac 60
ctgcancctt tggncactgc anattgaaac ctctcagttg ctgacatca cccatccctt 120
ggcgtggggt tccaccacaa ccactttgac tctgtggctc ctgnanggtg gntttctcgt 180
actggcagta tggaccttan ccnacatata cctctgttcc ctctgctnag anaaagaatt 240
cccttaacat gatataatcc acccatgcaa ntngctactg gccagctac cattttccat 300
ttgctcacag aatttcatte agtctacact ttggcattct ctctggcgat agagtgtggc 360
tgggctgacc gcaaaaggtg ccttacacac tggccccac cctcaaccgt tgacnctatc 420
gangcttgcc tctcctctt gattnncccc catgttggat atcagggtgc tcnagggtat 480
ggaaaagaaa caaaac 496

```

```

<210> 104
<211> 575
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc feature
<222> 18, 19, 45, 68, 77, 132, 155, 174, 219, 226, 238, 259, 263,
271, 273, 306, 323, 339, 363, 368, 370, 378, 381, 382, 436,
440, 449, 450, 456, 481, 485, 496, 503, 510, 512, 515, 528,
542, 552
<223> n = A,T,C or G

```

```

<400> 104
gcacctgctc tcaatccnnc tctcaccatg atcctccgcc tgcanaaact cctctgccaa 60
ctatggangt ggittcnggg gtggctcttg ccaactggga agaagccgtg gtgtctctac 120
ctgttcaact cngtttgtgt ctggggggtc aactnngggc tatggaagcg gctnaactgt 180
tgttttgttg gaagggctgg taattggctt tgggaagtng cttatnagaa ttggcctnng 240
gaagttgcta ttgaaagtng ccntggaagt ngntttgttg ggggtttttg ctggtggcct 300
ttgttnaatt ttgggtgcttt gtnaatggcg gccccctcnc ctgggcaatg aaaaaaatca 360
ccnatgcnng aaacctcnac nnaacagcct gggcttccct cacttcgaaa aaagtgtgct 420
ccccccaaa aaaggncaan ccctcaann ttggaangttg aaaaaatctc cgaatgggga 480
ncccnaaaac aaaaancccc ccntttcccn gnaanggggg aaataccncc ccccaactta 540
cnaaaacccct tntaaaaaac cccccgggaa aaaaa 575

```

```

<210> 105
<211> 619
<212> DNA
<213> Homo sapiens

```

```

<220>

```

<221> misc feature  
 <222> 260, 527, 560, 564, 566, 585, 599  
 <223> n = A,T,C or G

```
<400> 105
cactagttagg atagaaacac tgtgtccgga gagtaaggag agaagctact attgattaga 60
gcctaaaccca gggttaactgc aagaagaggc gggatacttt cagctttcca tgtaactgta 120
tgcataaagc caatgtagtc cagtttctaa gatcatgttc caagctaaact gaatcccact 180
tcaatacaca ctcatgaact cctgatggaa caataacagg cccaagcctg tggatgatg 240
tgcaactctg ctgactacta aaaaaatact actctcataa atgggtggga gtatttttgg 300
gacaaacctac tttgcttggc tgagtgaagg aatgatatto atatatctat ttattccatg 360
gacattttagt tagtgctttt tatataccag gcctgatgct gagtgacact ctgttgata 420
tttccaaatt tttgtacagt cgctgcacat atttgaatc atataaag acttccaaaa 480
aatgaagtcg ctggtttttc atggcaactt gatcagtaaa ggattcnctc ctgttttgta 540
cttaaaacat ctactatata gtttnanatga aattcctttt cccnctctcc cgaaaaaana 600
aagtgtggg gaaaaaaa 619
```

<210> 106  
 <211> 506  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc feature  
 <222> 8, 21, 31, 32, 58, 75, 89, 96, 99, 103, 122, 126, 147, 150,  
 158, 195, 210, 212, 219, 226, 246, 248, 249, 255, 258, 261,  
 263, 265, 275, 304, 317, 321, 331, 337, 340, 358, 371, 377,  
 380, 396, 450, 491  
 <223> n = A,T,C or G

```
<400> 106
cattgtgtnct ttcatttgc tntggaagtgt nnatctctaa cagtggacaa agttcccngt 60
gccttaaaact ctgtnaacct tttgggaant gaaaantng tantatgata ggttattctg 120
angtanagat gttctggata ccattanatn tgccccnngt gtcagaggct catatttctg 180
tatgtaaatg gtatntcatt cgctactatn antcaattng aaatanggtc tttgggttat 240
gaatantnng cagcncanct nanangctgt ctgtngtatt cattgtggtc atagcacctc 300
acancattgt aacctcnatc nagtgagaca nactagnaan ttctagtga tggtcanga 360
ttccaaatgg nctcatntcn aatgttttaa agttanttaa gtgtaagaaa tacagactgg 420
atgttcaccc aactagtcac tgtaatgaen ggctgtgcc aaacatctc ccttttccat 480
gactgtggta nccgcacatg gaaaaa 506
```

<210> 107  
 <211> 452  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc feature  
 <222> 289, 317, 378  
 <223> n = A,T,C or G

```
<400> 107
gttgagtctg tactaaacag taagatatct caatgaacca taaattcaac tttgtaaaaa 60
tcttttgaag catagataat attgttttgt aaatgtttct tttgttttgt aaatgtttct 120
tttaaaagacc ctccattctc ataaaactct gcactgtagg gctgtgttac ctttctctct 180
ctaagggtta caataggagt ggtgatttga aaaatatata attatgagat tgggttttcc 240
gtggcatata ttgcataact gtatcatttt cttttttaac cggtgaagant ttcaagttgt 300
tggaaggtaa ctgtganaac ccagttttcc gtccatctcc cttagggact acccatagaa 360
```

```
catgaaaagg tccccacnga agcaagaaga taagtctttc atggctgctg gttgcttaaa 420
ccactttaaa accaaaaaat tccccttgga aa 452
```

```
<210> 108
<211> 502
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<222> 22, 31, 126, 168, 183, 205, 219, 231, 236, 259, 283, 295,
296, 298, 301, 340, 354, 378, 383, 409, 433, 446, 455, 466,
488
<223> n = A,T,C or G
```

```
<400> 108
atcttcttcc cttaattagt tnttatttat ntattaaatt ttattgcatg ttoctggcaa 60
caaaaaagaga ttgtagattg gcttctggct ccccaaaagc ccataacaga aagtaccaca 120
agacnccaac tgaagcttaa aaaatctatc acatgtataa tacctttnga agaaccattaa 180
tanagcatat aaaaactttta acatntgctt aatgttgnac aattataaaa ntaatngaaa 240
aaaaatgtccc ttttaacatnc aatatcccac atagtgttat ttnaggggat taccnngnaa 300
naaaaaaagg gtagaaggga tttaatgaaa actctgcttn ccatttctgt ttanaaacgt 360
ctccagaaca aaaaacttntc aantctttca gctaacgcga ttgagctna ggccactcaa 420
aaactccatt agnccactt tctaangcto tctanagctt actaancett ttgacccett 480
accctggnata ctctgacct ca 502
```

```
<210> 109
<211> 1308
<212> DNA
<213> Homo sapiens
```

```
<400> 109
accgcaggctc tgcgtaaaat catcatggat tcacttggcg ccgtcagcac togacttggg 60
tttgatcttt toaaagagct gaagaaaaca aatgatggca acatctcttt ttoctctgtg 120
ggcatcttga ctgcaattgg catggtctct ctggggaccc gagggagcac cgcttccag 180
ttggaggagg tggtttcact tgaaaagagc acgaagagct caagaataaa ggctgaagaa 240
aaaagaggtga ttgagaacac agaagcagta catcaacatt tccaaaagtt tttagctgaa 300
ataagcaaac toactaatga ttatgaactg aacataacca acaggctggt tggagaaaaa 360
acataacctct tocttcaaaa atacttagat tatgttgaaa aatattatca tgcactctctg 420
gaacctgttg attttgttaa tgcagccgat gaaagtcgaa agaagattaa ttctctgggtt 480
gaaagcaaaa caaatgaaaa aatcaaggac ttgttccagc atggctctat tagtagctct 540
accgaagctgg tgcgtgtgaa catggtttat tttaaaggcg aatgggacag ggagttaaag 600
aaaagaaaata ctaaggaaag gaaattttg atgaataaga gcacaagtaa atctgtacag 660
atgatgacac agagccattc ctttagcttc actttctctg aggaacttga gcccaaaatt 720
ctagggtattc catataaaaa caacgaccta agcatgtttg tgcttctgcc caacgacato 780
gatggctctg agaagataat agataaaata agtctcgaga aattggtaga gtggactagt 840
ccagggcata tggaaagaag aaaggtgaat ctgcaacttg ccoggtttga ggtggaggac 900
agttacagat tagaggcgtt cctggtctgc atggggatgg gcgatgcttt cagtggagac 960
aaaagccgact actcgggaat gtcgtcagcg tccgggttgc acgccagcaa gttctgcac 1020
agttctcttg tggcagtaac tgaggaaggc accgaggtcg cagctgccac tggcataggc 1080
tttactgtca catccgcccc aggtcatgaa aatgttcaat gcaatcatcc ctctctctgt 1140
ttcatcagcg acaatgaatc caacagcact ctctctctgc cagatatttc ttctcttaa 1200
gatgatcggt gccatggcat tgctgctttt agcaaaaaac aactaccagt gttactcata 1260
tgattatgaa aatgctccat tcttttaaat ggtggctcac ttgcattt 1308
```

```
<210> 110
<211> 391
<212> PRT
```



&lt;213&gt; Homo sapiens

&lt;400&gt; 110

```

Met Asp Ser Leu Gly Ala Val Ser Thr Arg Leu Gly Phe Asp Leu Phe
 1          5          10
Lys Glu Leu Lys Lys Thr Asn Asp Gly Asn Ile Phe Phe Ser Pro Val
      20          25          30
Gly Ile Leu Thr Ala Ile Gly Met Val Leu Leu Gly Thr Arg Gly Ala
 35          40          45
Thr Ala Ser Gln Leu Glu Glu Val Phe His Ser Glu Lys Glu Thr Lys
 50          55          60
Ser Ser Arg Ile Lys Ala Glu Glu Lys Glu Val Ile Glu Asn Thr Glu
 65          70          75          80
Ala Val His Gln Gln Phe Gln Lys Phe Leu Thr Glu Ile Ser Lys Leu
      85          90          95
Thr Asn Asp Tyr Glu Leu Asn Ile Thr Asn Arg Leu Phe Gly Glu Lys
      100          105          110
Thr Tyr Leu Phe Leu Gln Lys Tyr Leu Asp Tyr Val Glu Lys Tyr Tyr
      115          120          125
His Ala Ser Leu Glu Pro Val Asp Phe Val Asn Ala Ala Asp Glu Ser
      130          135          140
Arg Lys Lys Ile Asn Ser Trp Val Glu Ser Lys Thr Asn Glu Lys Ile
      145          150          155          160
Lys Asp Leu Phe Pro Asp Gly Ser Ile Ser Ser Ser Thr Lys Leu Val
      165          170          175
Leu Val Asn Met Val Tyr Phe Lys Gly Gln Trp Asp Arg Glu Phe Lys
      180          185          190
Lys Glu Asn Thr Lys Glu Glu Lys Phe Trp Met Asn Lys Ser Thr Ser
      195          200          205
Lys Ser Val Gln Met Met Thr Gln Ser His Ser Phe Ser Phe Thr Phe
      210          215          220
Leu Glu Asp Leu Gln Ala Lys Ile Leu Gly Ile Pro Tyr Lys Asn Asn
      225          230          235          240
Asp Leu Ser Met Phe Val Leu Leu Pro Asn Asp Ile Asp Gly Leu Glu
      245          250          255
Lys Ile Ile Asp Lys Ile Ser Pro Glu Lys Leu Val Glu Trp Thr Ser
      260          265          270
Pro Gly His Met Glu Glu Arg Lys Val Asn Leu His Leu Pro Arg Phe
      275          280          285
Glu Val Glu Asp Ser Tyr Asp Leu Glu Ala Val Leu Ala Ala Met Gly
      290          295          300
Met Gly Asp Ala Phe Ser Glu His Lys Ala Asp Tyr Ser Gly Met Ser
      305          310          315          320
Ser Gly Ser Gly Leu Tyr Ala Gln Lys Phe Leu His Ser Ser Phe Val
      325          330          335
Ala Val Thr Glu Glu Gly Thr Glu Ala Ala Ala Thr Gly Ile Gly
      340          345          350
Phe Thr Val Thr Ser Ala Pro Gly His Glu Asn Val His Cys Asn His
      355          360          365
Pro Phe Leu Phe Phe Ile Arg His Asn Glu Ser Asn Ser Ile Leu Phe
      370          375          380
Phe Gly Arg Phe Ser Ser Pro
      385          390

```

&lt;210&gt; 111

&lt;211&gt; 1419

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 111

```

ggagaactat aaattaagga tcccagctac ttaattgact tatgtcttct agttcgttgc 60
ccagccacca cgtctctctc aaaaaccoga ggtctcgcta aaatcatcat ggattcactt 120
ggcgccgtca gcactcgact tgggtttgat cttttcaaag agctgagaa aacaatgat 180
ggcaacatct tcttttcccc tgtgggcate ttgactgcaa ttggcaggtt cctcctgtggg 240
acccgaggag ccaccggttc ccagttggag gaggtgtttc actctgaaaa agagacgag 300
agctcaagaa taaaggctga agaaaaagag gtggtaaaga taaaggctga aggaaaagag 360
attgagaaca cagaagcagt acatcaacaa ttccaaaagt ttttgactga aataagcaaa 420
ctcactaatg attatgaact gaacataacc aacaggctgt ttggagaaaa aacatacctc 480
ttccttcaaa aatacttaga ttatgttgaa aaatattatc atgcattctc ggaacctgtt 540
gattttgtaa atgcagccga tgaaagtoga aagaagatta attcctgggt tgaaagcaaa 600
acaaatgaaa aaatcaagga ctgttccca gatggctcta ttagtagctc taccaagctg 660
gtgctgtgtg acatgggtta ttttaaaggg caatgggaca gggagttaa gaaagaaaat 720
actaaggaag agaaattttg gatgaataag agcacaagta aatctgtaca gatgatgaca 780
cagagccatt cctttagctt cactttcctg gaggacttgc aggccaaaat tctagggatt 840
ccataaaaa caaacgacct aagcatgttt gtgcttctgc ccaacgacat cgtggcctg 900
gagaagataa tagataaaat aagtcctgag aaattggtag agtggaactag tccagggcatt 960
attggaagaaa gaaaggtgaa tctgcacttg ccccggttg agtggaagga cagtaccgat 1020
ctagaggcgg tcctggctgc catggggatg ggcgatgctc tcagtgcagc caaagccgac 1080
tactcgggaa tgcgtcagg ctccgggttg tacgccaga agttcctgca cagttccttt 1140
gtggcagtaa ctgaggagg caccgaggct gcagctgoca ttggcatagc ctcttactgtc 1200
acatccggcc caggtcatga aaatgttca cgcataatc ccttctgttt ctctcatagg 1260
cacaatgaat ccaacagcat cctcttcttc ggcagatttt ctctcctta agatgatcgt 1320
tgccatggca ttgctgcttt tagcaaaaa caactaccag tgttactcat atgattatga 1380
aaatcgtcca ttcttttaaa tgggtggctca cttgcattt 1419

```

&lt;210&gt; 112

&lt;211&gt; 400

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 112

```

Met Asp Ser Leu Gly Ala Val Ser Thr Arg Leu Gly Phe Asp Leu Phe
1      5      10      15
Lys Glu Leu Lys Lys Thr Asn Asp Gly Asn Ile Phe Phe Ser Pro Val
20      25      30
Gly Ile Leu Thr Ala Ile Gly Met Val Leu Leu Gly Thr Arg Gly Ala
35      40      45
Thr Ala Ser Gln Leu Glu Glu Val Phe His Ser Glu Lys Glu Thr Lys
50      55      60
Ser Ser Arg Ile Lys Ala Glu Glu Lys Glu Val Val Arg Ile Lys Ala
65      70      75      80
Glu Gly Lys Glu Ile Glu Asn Thr Glu Ala Val His Gln Gln Phe Gln
85      90      95
Lys Phe Leu Thr Glu Ile Ser Lys Leu Thr Asn Asp Tyr Glu Leu Asn
100     105     110
Ile Thr Asn Arg Leu Phe Gly Glu Lys Thr Tyr Leu Phe Leu Gln Lys
115     120     125
Tyr Leu Asp Tyr Val Glu Lys Tyr Tyr His Ala Ser Leu Glu Pro Val
130     135     140
Asp Phe Val Asn Ala Ala Asp Glu Ser Arg Lys Lys Ile Asn Ser Trp
145     150     155     160
Val Glu Ser Lys Thr Asn Glu Lys Ile Lys Asp Leu Phe Pro Asp Gly
165     170     175
Ser Ile Ser Ser Ser Thr Lys Leu Val Leu Val Asn Met Val Tyr Phe
180     185     190

```

Lys Gly Gln Trp Asp Arg Glu Phe Lys Lys Glu Asn Thr Lys Glu Glu  
 195 200 205  
 Lys Phe Trp Met Asn Lys Ser Thr Ser Lys Ser Val Gln Met Met Thr  
 210 215 220  
 Gln Ser His Ser Phe Ser Phe Thr Phe Leu Glu Asp Leu Gln Ala Lys  
 225 230 235 240  
 Ile Leu Gly Ile Pro Tyr Lys Asn Asn Asp Leu Ser Met Phe Val Leu  
 245 250 255  
 Leu Pro Asn Asp Ile Asp Gly Leu Glu Lys Ile Ile Asp Lys Ile Ser  
 260 265 270  
 Pro Glu Lys Leu Val Glu Trp Thr Ser Pro Gly His Met Glu Glu Arg  
 275 280 285  
 Lys Val Asn Leu His Leu Pro Arg Phe Glu Val Glu Asp Ser Tyr Asp  
 290 295 300  
 Leu Glu Ala Val Leu Ala Ala Met Gly Met Gly Asp Ala Phe Ser Glu  
 305 310 315 320  
 His Lys Ala Asp Tyr Ser Gly Met Ser Ser Gly Ser Gly Leu Tyr Ala  
 325 330 335  
 Gln Lys Phe Leu His Ser Ser Phe Val Ala Val Thr Glu Glu Gly Thr  
 340 345 350  
 Glu Ala Ala Ala Ala Thr Gly Ile Gly Phe Thr Val Thr Ser Ala Pro  
 355 360 365  
 Gly His Glu Asn Val His Cys Asn His Pro Phe Leu Phe Phe Ile Arg  
 370 375 380  
 His Asn Glu Ser Asn Ser Ile Leu Phe Phe Gly Arg Phe Ser Ser Pro  
 385 390 395 400

&lt;210&gt; 113

&lt;211&gt; 957

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 113

ctgcaccttc tctgcacagc ggatgaaccc tgagcagctg aagaccagaa aagccactat 60  
 gactttctgc ttaattcagc agcttacagg attcttcaaa gagtgtgtcc agcatccttt 120  
 gaaacatgag ttcttaccag cagaagcaga cctttacccc accactctag cttcaacagc 180  
 agcagggtgaa acaaccagc cagcctccac ctcaggaaat atttgttccc acaaccaagg 240  
 agccatgccca ctcaaaaggtt ccacaacctg gaaacacaaa gattccagag ccaggctgta 300  
 ccaaggtccc tgagccaggc tgtaccaagg tccctgagcc aggttgtacc aaggtccctg 360  
 agccaggatg taccaaggtc cctgagccag gttgtaccaa ggctccctgag ccaggctaca 420  
 ccaaggtccc tgagccaggc agcatcaagg tccctgacca aggttctcat aagtttctg 480  
 agccaggatg catcaaagtt cctgagcaag gatacaccaa agttcctgtg ccaggctaca 540  
 caaaggttacc agagccatgt ccttcaacgg tcaactccagg cccagctcag cagaagacca 600  
 agcagaagta atttgtgtga cagacaagcc ctgagaagc caaccaaccag atgtgggaca 660  
 cccctctccc atctgtttct gtgtcttaat tgtctgtaga ccttgttaac agtacattct 720  
 caccocaaagc catagctctc ctcttatttg tatcctaaaa atacggctact ataaagcttt 780  
 tgttcacaca cactctgaag aatcctgtaa gccctgaaat taagcagaaa gtcttcatgg 840  
 ctttctgggt cttcggtgct tcagggttca tctgaagatt cgaatgaaaa gaaatgcagt 900  
 tttctgctgc tgcctcatt aaattgcttt taattccaaa aaaaaaaa 957

&lt;210&gt; 114

&lt;211&gt; 161

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 114

Met Ser Ser Tyr Gln Gln Lys Gln Thr Phe Thr Pro Pro Pro Gln Leu

1	5	10	15
Gln Gln Gln Gln Val	Lys Gln Pro Ser	Gln Pro Pro Pro	Gln Glu Ile
20	25	30	
Phe Val Pro Thr Thr	Lys Glu Pro Cys	His Ser Lys Val	Pro Gln Pro
35	40	45	
Gly Asn Thr Lys Ile	Pro Glu Pro Gly	Cys Thr Lys Val	Pro Glu Pro
50	55	60	
Gly Cys Thr Lys Val	Pro Glu Pro Gly	Cys Thr Lys Val	Pro Glu Pro
65	70	75	80
Gly Cys Thr Lys Val	Pro Glu Pro Gly	Cys Thr Lys Val	Pro Glu Pro
85	90	95	
Gly Tyr Thr Lys Val	Pro Glu Pro Gly	Ser Ile Lys Val	Pro Asp Gln
100	105	110	
Gly Phe Ile Lys Phe	Pro Glu Pro Gly	Ala Ile Lys Val	Pro Glu Gln
115	120	125	
Gly Tyr Thr Lys Val	Pro Val Pro Gly	Tyr Thr Lys Val	Pro Glu Pro
130	135	140	
Cys Pro Ser Thr Val	Thr Pro Gly Pro	Ala Gln Lys	Thr Lys Gln
145	150	155	160
Lys			

&lt;210&gt; 115

&lt;211&gt; 506

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

<222> 8, 21, 31, 32, 58, 75, 89, 96, 99, 103, 122, 126, 147, 150,  
 158, 195, 210, 212, 219, 226, 246, 248, 249, 255, 258, 261,  
 263, 265, 275, 304, 317, 321, 331, 337, 340, 358, 371, 377,  
 380, 396, 450, 491

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 115

cattggtnc	ttcatttgc	ntggaagtgt	nnatctctaa	cagtggaaca	agttccnct	60
gccttaaact	ctgtnacact	tttgggaant	gaaaanttng	tantatgata	ggttattctg	120
angtanagat	gttctggata	ccattanatin	tgccccnctg	gtcagaggct	catattgtgt	180
tatgtaaatg	gtatntcatt	cgctactatn	antcaattng	aaatanggtc	tttgggttat	240
gaatantnng	cagcncanct	nanangctgt	ctgtngtatt	cattgtgtgc	atagcacctc	300
acancattgt	aacctonact	naagttagaca	nactagnaan	ttcctagtga	tggctcanga	360
ttccaaatgg	nctcatntcn	aatgttttaa	agttanttaa	gtgtaagaaa	tacagactgg	420
atgttccacc	aactagtacc	tgtaatgacn	ggcctgtccc	aacacatctc	cettttccat	480
gactgtggta	ncccgcatcg	gaaaaa				506

&lt;210&gt; 116

&lt;211&gt; 3079

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 116

ggatccccgg	gtttcctaaa	ccccccacag	agtcctgcc	aggccaaaga	gcaaggaaaa	60
ggatcaagg	cagaaaaaat	gctgagttag	gaggagctat	ggaaggataa	acctggcctt	120
aaagaggtca	aagtgtttta	tagggggcgc	tgagggtctc	ccacattctc	tggcctaaac	180
cttgaggcca	gatctgccca	gtgggcctcg	ggatagctgt	gccttcoccta	acaaaaaaat	240
tgtagcacia	aggatgaaac	tctattttcc	ctctagcaca	taaccaagaa	tataaggcta	300

cagattgacct tccccagagg gaaaaccctg cagcaacctg ctgctgtgaa aagtgtgaaga 360  
 gcagatcaact ggggaatcgt ttgccccccg ctgatggaca gcttccccaa gctccaaaggg 420  
 cagggtgctca gcatgtacog tactgggatg gttgtcaata ctctgtgtcc tgtaagagatc 480  
 ccaggacact gccatgccaa tgccccctca gttctctgga tccttttttg gctgtccaca 540  
 gccccagcct ctatggtgaa gacataactg ctgacagcgt caccaaactg ttgccaagag 600  
 atcagtgtct gaaggcaagg ttattttctaa ctgagcagag cctgcacagg agcaagaggt 660  
 tgcacccccc accactgtgc aggtgtgacc ggtgagctca cagctgcctcc cagggcagtc 720  
 ccagccctcc taatcatcac agctcgacag ctctctcgcc cagccagagt ctggaaggga 780  
 taaaaagggg catcacccgt cctgggtaac agagccacct tctgogtctc gctgagctct 840  
 gttctctcca gcaactccca accoactagt gcoctggtct ctgtgtccac caggaaacaag 900  
 ccacoatgtc tgccagatca agtgtgtctt ccggagcggg gggcagtcgt agcttcagca 960  
 ccgctctcgc catcaccccg tctgtctccc gacacagctt caoctcoctg tcccggtccg 1020  
 ggggtggcgg tgggtggtgc ttccggcaggg tcagccttgc ggtgtctgtt ggagtggggt 1080  
 gctatggcag ccggagcctc tacaacctgg ggggtcccaa gaggatatcc atcagcacta 1140  
 gtggtggcag cttcaggaaac cggtttgggt ctggtgtcgg aggcggctat ggccttggag 1200  
 gtggtgcagg tagtggtatt ggtttccggg gttgagctgg ttgtgtgttt ggcctcgggt 1260  
 gcggagctgg ctttggaggt ggcctcgggt gcctgtctgc cctcctggag 1320  
 gttccaaaga ggtcactgtc aaccagagtc tctgaactcc cctcaactgc caaatcgacc 1380  
 ccagcatcca gaggtgagg accgaggagc gcgagcagat caagaccctc aacaataagt 1440  
 ttgctcctct catcgacaag gtgcggttcc tggagcagca gaacaaggtt ctggaacaaa 1500  
 agtggacctc gctgcaggag cagggcacca agactgtgag gcagaaacct gagccgttgt 1560  
 tcgagcagta catcaacaac ctcaggaggg agctggacag catctggggg gaacggggcc 1620  
 gcctggactc agagctgaga aacatgcagg acctggtgga agacttcaag aacaagtatg 1680  
 aggatgaat caacaagcgt accactgctg agaattgagt tgtgatctgc aagaaggatg 1740  
 tagatgctgc ctacatgaac aaggtggagc tggaggccaa gtttgatgca ctgatggatg 1800  
 agattaaact catgaagatg ttctttgatg cggagctgtc ccagatgcag acgatgtct 1860  
 ctgacacctc agtggctctc tccatggaca acaaccggaa cctggacctg gatagcatca 1920  
 tcgctgaggt caagggcccg tatgagaga ttgccaaacg cagccggaca gcgcgcagat 1980  
 cctggtatca gaccaagatg gaggagctgc agcagacagc tggccggcat ggcgatgacc 2040  
 tcggcaacac caagcatgag atctctgaga tgaacagcag gatccagagc ctgagagcct 2100  
 agattgacaa tgtcaagaaa cagtgcgccca atctgcagaa cgcattgccc gatcccgagc 2160  
 agcgtgggga gctggccctc aaggatgccca ggaacaagct ggccgagctg gaggaggccc 2220  
 tgcagaaagg caagcaggac atggcccgcc tctcggtgag gtaccaggag ctoatgaaca 2280  
 ccaagctggc cctggacgtg gagatgccca cttacggcaa gctgtggag ggcgaggaaat 2340  
 gcagactcag tggagaaggg gttggaccag tcaacatctc ttgtgtcaca agcagtggtt 2400  
 cctctggata tggcagtgcc agtggtctat gcggtggcct cgggtggaggt ctggggcgcc 2460  
 gcctcgggtg aggtcttgcc ggaggtagca gtgaaagcta ctactccagc agcagtgagg 2520  
 gtgtcggcct aggtggtggg ctcagtggtg ggggtctctg cttcagtgca agcagtagcc 2580  
 gagggtcggg ggtgggcttt ggcagtgggg ggggtagcag ctccacgcgc aaatttgtct 2640  
 ccaccaactc ctctcccgcg aagagcttoa acctgctgca agctactgoc gtaactgoc 2700  
 ttccaagctg agcaacccag cccatggaga ttgcctcttc taggcagttg ctcaagccat 2760  
 gttttatcct tttctggaga ctagtctaga ccaagccaat tgcagaaacca cattctttgg 2820  
 ttccacaggag agccccaatt ccagccctcg gtctcccgct cgcagttct atattctgct 2880  
 tcaaatcagc cttcaggttt cccacagcat ggccctgtct gacaogagaa cccaagttt 2940  
 tcccaaatct aatcatcaca aacagaatoc ccaccccaat cccaaatttt gttttggtc 3000  
 taactaacct cagaatgtgt tcaataaaat gttttataat ataagctggt gtgcagaatt 3060  
 gttttttttt tcatcccaa 3079

&lt;210&gt; 117

&lt;211&gt; 6921

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 117

gaattctgac tgtocactca aaacttctat tccgatcaaa gctatctgtg actacagaca 60  
 aattgagata accatttata agacagatga atgtgttttg gcgaataact ctcatcgtgc 120  
 taaatggag gtacttagtc ctactgggaa tgaggctatg tgcocactgt ctcatctcac 180  
 cgttctctcca ccaacaaaag aagcgggtgga ccttgccaac agaattgagc aacagtatca 240

gaatgtcctg	actctttggc	atgagttctc	cataaacatg	aagagtgtag	tatcctggca	300
ttatctcctc	aatgaaattg	atagaattcg	agctagcaat	gtggcttcaa	taaaagacaa	360
gctacctgat	gaacctcagc	aagttctaa	taactctacaa	tctcgttttg	aagattttct	420
ggaagatagc	caggaatccc	aagttctttc	aggctcagat	atacaacaac	tggaaaagga	480
ggtttaattg	tgtaagcagt	attatcaaga	acttctttaa	tctgcagaaa	gagaggagca	540
agaggaatca	gttttaatac	tctacatctc	tgaggttcga	aacattagac	tccggttaga	600
gaaactgtgaa	gatcggtgtg	ttagacagat	togaactccc	ctggaagagc	atgattttgca	660
tgaagtggtg	gagaacatca	cagaacagga	gaaactaaag	aaagagctgg	aacgacttaa	720
agatgatttg	ggaacaatca	caaataagtg	tgaggagttt	ttcagtcagc	cagcagcctc	780
ttcatcagtc	cctaccctac	gatcagagct	taagtgtgtc	cttcagaaac	tgaaccaaag	840
ctattctcat	tcttccactt	acatagataa	gttgaaaact	gttaactctg	tgttaaaaaa	900
cactcaagct	gcagaagccc	togtataaac	ctatgaaact	aaactgtgtg	aagaagaagc	960
agttatagct	gacaagaata	atattgagaa	tctaataagt	actttaaagc	aatggagatc	1020
tgaagttagt	gaaaagagac	aggtattcca	tgctcttaga	gatgagttgc	agaaaagctaa	1080
agccatcag	gatgaaatgt	ttaaaacgta	taaagaacgg	gacctgtatt	ttgactggca	1140
acagaagaaa	gcagatcaat	tagttgaaag	gtggcaaaat	gttcatgtgc	agattgacaa	1200
caggtttacg	gacttagagg	gcattggcaa	atcactgaag	tactacagag	acacttacca	1260
tctttagat	gatttgatcc	agcaggttga	aactactcag	agaaagattc	aggaataatc	1320
gcctgaaatg	agtaaaaaac	tagccacaca	gttgaatcaa	cagaagatgc	tggtgtccga	1380
aataagaatg	aaacagagca	aaatggacga	gtgtcaaaaa	tatgcagaac	agactctcagc	1440
tacagtgaag	gactatgaat	tacaacaact	gacctaccgg	gccatggtag	attcacaaca	1500
aaaaatccca	gtgaaacgag	gaagaatgca	gagttcagca	gatctcaata	ttcaagaagt	1560
catggaacct	aggactcgat	atactgcoct	ggtcactctc	atgcacacaa	atattaagatt	1620
tgtcgtgtat	tcatttgaaga	ggctggaaga	ggaggagatt	aaaaggtgta	aggagacttc	1680
tgaacatggg	gcataattcag	atctgcttca	gcgtcagaag	gcaacagtcg	ttgagaatag	1740
caaaacttaca	ggaaaagata	gtgagttgga	aagaatggta	gctgaactaa	agaaaacaaa	1800
gtcccgagta	gaggaagaac	ttccgaaggt	cagggaggct	gcagaaaaat	aattgagaaa	1860
gcagcagaga	aagtgaaga	atatctctct	gcagaaagata	agggctgaaa	gaatggacaa	1920
gcagtaaccg	agggaaacttg	aaaccattgt	gagagagaag	gaagccgctg	aaaagaaact	1980
ggagcgggtg	aggcagctca	ccatagaggc	cagaggctaaa	agagctccgc	tggaaagaga	2040
cctcctgaat	tttcgcaatc	agttggaggga	aaacaccttt	accagacgaa	catcgtgaag	2100
tcactctaaa	agaaaagatt	taagtctcaa	tgatttggag	caacaaaaaa	ataaattaat	2160
ggaagaatta	agaagaaga	gagacaatga	ggaagaactc	ttgaagctga	taaagcagat	2220
ggaaaaagac	cttgcatctc	agaaacaggt	agcagagaaa	cagtttgaag	aaaagcagaa	2280
aattgaattg	gaagcagaag	gaaaaataac	tgaatttcag	tatacatgta	gagaaaaatgc	2340
attgccagtg	tgtccgatac	cacaggctac	atcatgcagg	gcagtaacgc	gtctccagca	2400
agaaactgac	aagcagaagc	cagaagaact	caaacagcag	ctagatgaac	taacagctgc	2460
caatagaaag	gctgaacaa	acatgagaga	gctgacatat	gaacttaagt	ccctccagct	2520
tgaaaaaacg	tcatctgagg	aaaaggctcg	tttgctaaaa	gataaactag	atgaaacaaa	2580
taatacactc	tgactgctta	agttggagct	ggaagggaag	gatcaggcgg	agaaagggta	2640
ttctcaacaa	ctcagagagc	tttgtaggca	attgaatcaa	accacaggta	aagctgaaga	2700
agccatgcga	gaagcttagt	atctcaagaa	aatataagct	aatattcagt	tgaattaga	2760
atctcttaat	catgaaaaag	ggaactacaa	aagagaagta	gcaggaatca	caaagggaaca	2820
tgcgttagct	gagaagaata	ttcagcattt	aaattcaca	tatctctctt	tctgagatga	2880
gaaagaatta	gaaagactac	aaatctgcca	gagaaaaatc	gatcatctaa	aagaacaatt	2940
tgaagaagaac	catgagcagt	tgcttcaaaa	tatcaaaagt	gaaaaagaaa	ataatgataa	3000
aatccaaagg	ctcaatgaag	aattggagaa	aagtaatgag	tgtgcagaga	tgtcaaaaac	3060
aaaagttaga	gagcttacta	ggcagaataa	tgaacccaaa	ttaatgatgc	agagaattca	3120
ggcagaatga	gagaattata	ttttagagaa	acaaactatc	cagcaaaagt	gtgaagactc	3180
gaaaattcag	gcagatggtt	ttaaagatca	gctacgcagc	acaaatgaac	acttgcataa	3240
acagacaaaa	acagagcagg	attttcaag	aaaaattaaa	tgctctgaag	aagacctggc	3300
gaaaagtcaa	aatttggtaa	gtgaatttaa	gcaaaagtgt	gaccaacaga	acattatgat	3360
ccagaatacc	aagaagaagc	ttgaaaattc	gaatcgagg	ctgaatgcgt	ccaaagaaga	3420
gaagcgaagc	ggggagcaga	aagttcagct	acaaacagct	cgaggtcaag	gttaataata	3480
caggttgaaa	aaagtacaa	acgaattaca	cttaaaagac	acagagcagc	agatgaccca	3540
cagaagaatg	gttctgtttc	aggaagaatc	tggttaattc	ataaagcagc	cagaggagtt	3600
tcggaagaag	atggaaaaat	taatggagtc	caaagtcatc	actgaaaagt	atatctcagg	3660
cattaggctt	gactttggtg	ctcttcaaca	agaaaactct	agagcccaag	aaaatgctaa	3720

gctttgtgaa	acaaacatta	aagaacttga	aagacagctt	caacagtatc	gtgaacaaat	3780
gcagcaagg	cagcacatgg	aagcaaatca	ttaccaaaaa	tgtcagaaac	ttgaggatga	3840
gctgatagcc	cagaagcgtg	aggttgaaaa	cctgaagcaa	aaaatggacc	aacagatcaa	3900
agagcatgaa	catcaaatag	ttttgctcca	gtgtgaaatt	caaaaaaaga	gcacagccaa	3960
agactgtacc	ttcaaacccag	attttgagat	gacagtgaag	gagtgccacg	actctgggaa	4020
gctgtcctct	agaaacactg	gacaccttca	cccaacaccc	agatcccttc	tgttgagatg	4080
gactcaagaa	ccacagccat	tggaagagaa	gtggcagcat	cgggttgttg	aacagatacc	4140
caaaagagtc	caattccagc	caccaggggc	tcactctgag	aaagagaaaa	gccacagctg	4200
ttactctgag	tactttttct	agacaagcac	cgagttacag	ataacttttg	attgagacaaa	4260
ccccattaca	agactgtctg	aaattgagaa	gataagagac	caagccctga	acaatttctag	4320
accocctgtt	aggtatcaag	ataacgcctg	tgaaatggaa	ctggtgaagg	ttttgacacc	4380
cttagagata	gctaagaaca	agcagtatga	tatgcataca	gaagtcacaa	catataaaca	4440
agaaaaagac	ccagttccca	gtgctgaaga	atggatgctt	gaagggtgca	gagcatctgg	4500
tggaactcaag	aaaggggatt	tccttaagaa	gggcttagaa	ccagagacct	tcacagaactt	4560
tgatgggtgat	catgcactgtt	cagtcaggga	tgatgaattt	aaattccaag	ggcattaggca	4620
cactgtgtgact	gccagcgagct	tggtggaaagc	taagctttctg	gcacatgagaa	caattgagca	4680
gctgcagctac	ggctttaaga	ctgttgaaga	agttcagaaa	actcttaaca	agttctctgac	4740
gaaagccacc	tcaattggag	ggctttacct	agaactctaca	aaagaaaaga	tttcaatttg	4800
ctcagcgggcc	gagagaaatca	taatatagaaa	aatggtggct	ttggcatttt	tagaagctca	4860
ggctgcaaca	ggttttataa	ttgatcccat	ttcaggtcag	acatatcttg	ttgaagatgc	4920
agttctctaaa	ggctgtgttg	accccgaatt	cagaattagg	cgattttgag	caattgagca	4980
agctgtgggga	tattcttatt	cttctaaagc	attgtcagtg	tttcaagcta	tggaaaaatag	5040
aatgctgtgac	agacaaaaag	gtaaaacatat	cttggaaagc	cagatttgcca	gtggggtgtg	5100
caattgacccct	gtgagaggca	ttcgtgttcc	tcagaaaatt	gctctgcagc	aggggtttgt	5160
gaataatgcc	atctttacagt	ttttacatga	gccatccagc	aacacaaag	ttttccotaa	5220
toccaataac	aagcaagctc	tgtattactc	agaattactg	cgaatgtgtg	tattttagatg	5280
agagtcoccaa	tgctttctgt	ttccattttg	ggagaggaac	atttccaatc	ccaattgtcaa	5340
gaaaaacacat	agaattttctg	tagtagatac	taaaaacagga	tcagaattga	ccgtgtatga	5400
ggctttccag	agaaaacctga	ttgagaaaag	tatatatctt	gaactttcag	ggcagcaata	5460
tcagtggaag	gaagctatgt	tttttgaatc	ctatgggcat	ttcttctata	tctgtactga	5520
tactaaaaca	ggatttacact	tcaatatata	tgaggctata	gagcagggaa	caattgacaa	5580
agccttggtc	aaaaagtatc	aggaaggcct	catcacactt	acagaacttg	ctgattctctt	5640
ctgtgagccg	ttagtcccca	agaaaagattc	gcacagtcct	gttgccagggt	atttggctgac	5700
tgctagtggg	gaaaggatct	ctgtactaaa	agcctccctg	agaaatttgg	ttgatccggt	5760
tactgcctct	cgatgccttg	aagcccaagt	cagtcacagg	ggcataattg	atcctcttcc	5820
tggaacaaaag	taccgggttg	ccgaagcttt	gcacataggc	ctgggttagat	aggggtttgc	5880
ccagcagctg	cgacagtgtg	aattagtaat	caacagggtt	ggccatccca	tcactaacaa	5940
aatgatgtca	gtggtggaag	ctgtgaatgc	aaatatata	aataaggaaa	tgggaatccg	6000
atgtttggaa	tttcatgact	tgacaggagg	gttgatagag	ccacagggtt	actctcggtt	6060
atcaatagaa	gaggctctcc	aagttagtat	tatagatgct	ctcatttgcca	caaaactcaa	6120
agatcaaaaag	ctcatatgtca	gaaatataat	atgccctcag	acaaaagaaa	agttgacata	6180
taaaagaagcc	ttagaaaaag	ctgattttga	tttccacaca	ggaacttaaac	gcttagaagt	6240
atctgagccc	ctgatgacag	gaattttctag	cctctactat	tcttctaat	gggacatgtt	6300
taaaataact	tgcaagggtg	gatgcaggct	ggttcatgcc	actttttcag	agtatgatga	6360
tatcggctgt	atatgcagtc	tgtgaattat	gtaacatact	ctattttctg	aggggtgcga	6420
attgctaagt	gctcaaaata	gagtaagttt	tgaattgaaa	attacataag	atttataagg	6480
cttcaaatgt	tttcattttag	cottgagaat	ggttttttga	acattggcca	cactaaatag	6540
tttttttttt	tttaactaga	atgtgggata	aacttgatga	actccaagtt	cacagtgctca	6600
tttcttcaga	actccctctc	attgaatagt	gatcatttat	taaatgataa	attgcactcg	6660
ctgaaagagc	acgtcatgaa	gcaccatgga	atcaaaagaga	aagatatata	tcgttccoca	6720
cagccttcaa	gctgcagtg	tttagattgc	ttcaaaaaat	gaaaaagatt	tgccctttttc	6780
gatatagtga	cctcttttgc	atattaaaat	gtttaccaca	attgcacatt	tctgatttaag	6840
tcttgcacat	tgaagctcaa	cattatgaat	attatgtgtt	ggagaggggg	aaggattttc	6900
ttcattctgt	gtattttccg	g				6921

<211> 946  
 <212> DNA  
 <213> Homo sapiens

<400> 118

cttctgactg	ggctcaggct	gacaggtaga	gctcaccatg	gottcttctg	tccttctccc	60
ctccccatca	cagctgtggg	gcagtcacac	gtctccagtg	gctatggcgg	tgccagtggt	120
gtcggcagtg	cttaggcct	gggtggagga	agcagctatg	cctatggcag	tggtcttggc	180
gttgagggtg	gcttcagttc	cagcagtgga	agagccattg	gggggtggcc	cagctctgtt	240
ggaggcgcca	gttccaccat	caagtacacc	accacctcct	cctccagcag	gaagagctat	300
aagcactaaa	gtgcgtctgc	tagctctcgg	tcccacagtc	ctcaggccccc	ttcttggctg	360
cagagccctc	tcctcaggtt	gootgtcctc	tcttggtcct	cagctctccc	tgctgtccca	420
ggtagagctg	gggatgaatg	cttagtgccc	tcacttcttc	ttctctcttc	tataccatct	480
gagcaccocat	tgctcaccat	cagatcaacc	tctgatttta	catcatgatg	taatcaccac	540
tggagcttca	ctgttactaa	attattaatt	tcttgctccc	agtgtttctat	ctctgaggct	600
gagcattata	agaaaatgac	ctctgctcct	tttcatttga	gaaaattggc	aggggcttat	660
ttcagaacaa	cttccactta	ctttccactg	gctctcaaac	ttcttaacct	ataagtgttg	720
tgaaccccca	cccaggcagt	atccatgaaa	gcacaagtga	ctagtctcat	gatgtacaaa	780
gcctgtatct	ctgtgatgat	ttctgtgctc	ttcaactgtt	gcaattgcta	aataaagcag	840
atttataata	catataattc	tttacttttg	cttgcttttg	ggccaaagtt	ttgggcttaa	900
acttttttat	ctgataagtg	aatagttgtt	ttttaaagat	aatcta		946

<210> 119  
 <211> 8948  
 <212> DNA  
 <213> Homo sapiens

<400> 119

taaacagccc	ctgctccttg	ggccccctca	tgccatgcgc	taatctctcc	cacccgacca	60
acaccaacac	ccagctccga	cgcagctcct	ctgcgccttc	gcgcgcctcc	gagccacagc	120
tttctctccc	ctcctgcccc	cgcccccctg	cgtctctccc	gtcgcagcgc	gctctgggag	180
ggcccaggta	gcgagcagcg	acctcgcgag	ccttcgcacg	ttcccgcctg	ttccccggcc	240
gtccgcctat	ccttggcccc	ctccgctttc	tcgcgcctgc	ccgcctctgc	ttatgcctcg	300
gcgtgagcc	gctctccaga	ttgcccgcgc	acatgagctg	caacggaggc	ttccaccctg	360
ggatcaaac	tctggggcgc	atgatccgcg	ccgagctctg	cccggaacct	cgctacaggc	420
tgaccagcgc	cgccgggggc	accagcagga	tgtactattc	tcggcgccgc	gtgatccacc	480
accagaagct	ggagccctac	tgtcaaacgc	gcacgatctc	caggcacaag	aaccagaaca	540
ccatccagga	gctgctgcag	aactgctccg	actgcttgat	gcgagcagag	ctcatctgtc	600
agcctgaatt	gaagtatgga	gatggaatac	aactgactcg	gagtcagaaa	ttggatagat	660
gttttgcoca	ggccaatgac	caaatggaaa	ttctgcagac	cttgatcaga	gagatgcggc	720
agatggggca	gccctgtgat	gcttaccaga	aaaggctctt	tcagctccaa	gagcaaatgc	780
gagcccttta	taaaagccatc	agtgctccctc	gagtcgcgag	ggccagctcc	aagggtgggt	840
gaggctacac	ttgtcagagt	ggctctggct	gggatgagtt	caccaaacat	gtcaccagtg	900
aatgtttggg	gtggatgagg	cagcaaaagg	cggagatgga	catggtggcc	tggggtgtgg	960
acctggcttc	agtggagcag	cacattcaac	gccaccgggg	catccaccca	ttcatcctga	1020
actatcgctg	gcagctggac	aaaatcaaa	ccgacctcgc	cgagaacctt	gcgacttacc	1080
agtttgagga	ggagtatgaa	aaactgctga	aagctctctt	tgagaggatg	gatcacctgc	1140
gacagctgca	gaacatcat	caggccacgt	ccaggagagt	catgtggatc	aatgaactgc	1200
aggagaggga	gctgctgtac	gactggagcg	acaagaacac	caacatcgct	cagaacacag	1260
aggcctcttc	catacgcgat	agtcacactg	aagttaaaga	aaaagagctc	ataaagctga	1320
aacaagaag	tgaccaactc	gtctccaatc	agcatccagc	ttcagacaaa	attgaggcct	1380
atatggacac	ttctgcagacg	cagtgtagtt	ggattcttca	gatccacca	tgcatgtgat	1440
ttcatctgaa	agaaaatgct	gcttactttc	agttttttga	agaggcgacg	tcatctgaag	1500
catacctgaa	ggggctccag	gactccatca	ggaagaagta	ccctcgcgac	aaacaacatc	1560
ccctgcagca	cctgctggaa	cagatcaagg	agctggagaa	agaacagag	aaaatccttg	1620
aatacaagcg	tcaggtgacg	aacttggtta	acaagtcctaa	gaagattgta	cagctgaagc	1680
ctcgttaacc	agaactacaga	agcaataaac	ccattattct	cagagctctc	ttgtactaca	1740
aacaagatca	gaaaatcgtg	cataaggggg	atgagtgat	cctgaaggac	aacaacagac	1800



gcagcaagt	gtacgtgaag	ggcccgagg	gcgttgacat	gcttgttccc	tctgtggggc	1860
tgatcatccc	tccctcgaaac	ccactggcgc	tggaacctctc	ttgcgaagatt	gagcagtaact	1920
acgaagccat	cttggctctcg	tggaaccagc	tctacatcaa	catgaagagc	ctgggtgtcc	1980
ggcactactg	catgattgac	atagagaaga	tcaggggccat	gacaatcgcc	aagctgaaaa	2040
caatgocgca	ggaagatttac	atgaagacga	tagccgacct	tgagttacat	taccaagagt	2100
tcacagaaaa	tagccaaggc	tcagagatgt	ttggagatga	tgacaagcgg	aaaaacaggt	2160
ctcagtttgc	cgatgcccgac	aagcattacc	agaccctggt	cattcagctc	cctggctatc	2220
cccgacccca	gacagtgacc	acaactgaaa	tcactcatca	tggaacctgc	caagatgtca	2280
accataataa	agtaattgaa	accaacagag	aaaatgacaa	gcaagaacaa	tgatgtctga	2340
tgagctgcga	gaagattcgc	aggcagatag	agcactgcga	gggcagggatg	actctcaaaa	2400
acctccctct	agcgagccag	gggtcttctc	accacatcac	agtgaaaatt	aacgagctta	2460
agagtgctga	gaatgattca	caagcaattg	ctgaggttct	caaccagctt	aaagatattg	2520
ttgccaaact	cagaggttct	gaaaagtact	gctattttaca	gaatgaagta	tttggaactat	2580
ttcagaaact	ggaaaatatc	aatggtgtta	cagatggcta	cttaaatagc	ttatgcacag	2640
tgatgggcat	gctccaggct	attctccaaa	cagaagacat	gttaaagggt	tatgaagcca	2700
gaactcatga	ggaggaaact	gtctgcctgg	acctggatac	agtgcgaagct	taccgctgtg	2760
gactgaagaa	aataaaaaat	gacttgaact	tgaagaagtc	gttgtgtgcc	actatgaaga	2820
cagaaactca	gaagcccgag	cagatccaat	ctcagacttc	acgacagctc	ccactttatg	2880
atctggactt	gggcaagttg	gggtgaaaaa	tcacacagct	gacagaccgc	tgccaaagga	2940
tagataaaca	gatcgacttt	agattatggg	acctggagaa	acaaactcaag	caattgagga	3000
attatcgtga	taactatcag	gctttctgca	agtggctcta	tgatcgtaaa	cgccgccagg	3060
atctcctaga	atccatgaaa	tttggagatt	ccaacacagt	catgcggttt	ttgaatgagc	3120
agaaactaga	gcacagtgaa	atatctggca	aacgagacaa	atcagagaaa	gtacaaaaaa	3180
ttgtcgaaat	ttgcgcgaat	tcattaaagg	attatgagct	ccagctggcc	tcataccact	3240
caggagctga	aactctgctg	aacataccta	tcaagaggac	catgattcag	tcocctctcg	3300
gggtgattct	gcaagaggct	gcagatgttc	atgctcggtt	cattgaaacta	cttacaagat	3360
ctggagacta	ttacaggttc	ttaagtgaga	tgctgaagag	tttggaaagt	ctgaagctga	3420
aaaaatccaa	gatcgagatt	ttggaagagg	agctcagact	ggcccgagag	gcacactcgg	3480
aaaaactgtaa	taagaaacaa	ttctcggtac	agaacctgca	gaaataccag	gcagagtggt	3540
ccagcttcaa	gtcgcaactt	gcgagcctgg	aggagctgaa	gagacaggct	gagctgtgatg	3600
ggaagtgcgc	taagcaaaat	ctagacaagt	gctacggcca	aataaaaaga	ctcaatgaga	3660
agatcaaccg	actgacttat	gagattgaag	atgaaaagag	aagaagaaaa	tcctgtggaag	3720
acagatttga	ccaacagaa	aatgactatg	accaactgca	gaagcagaaa	caatgtgaaa	3780
aggagaacct	tggtttggcag	aaattagagt	ctgagaagag	catcaaggag	aaggagtacg	3840
agattgaaag	gttgagggtt	ctactgcagg	aagaaggcac	ccggaagaga	gaatatgaaa	3900
atgagctggc	aaaggtaaga	aaccactata	atgaggagat	gagtaaatga	aggaacaagt	3960
atgaaaacga	gattacaact	acgaagacca	ccatcaaggca	gatatccatt	caaaagaggg	4020
atgattccaa	aaatcttaga	aaccagcttg	atagactttc	aagggaaaaa	cgagatctga	4080
aggatgaaat	tgtaggcttc	aatgacagca	tcttgtaggc	cactgagcag	cgaaggcgag	4140
ctgaagaaaa	cgccctctag	caaaaggcct	gtggctctga	gataatgcag	aagaagcagc	4200
atctggagat	agaaactgaag	caggctcatgc	agcagcgctc	tgaggacaat	gcccggcaca	4260
agcagctccg	ggaggaggct	gccaaagcca	ttcaggacaa	aaataaggag	atcgagagac	4320
tcaaagctga	gtttcaggag	gaggccaagc	gccgctggga	atatgaaat	gaactgagta	4380
aggtaaagaa	caattatgat	gaggagatca	ttagcttaaa	aaatcagttt	gagacccaga	4440
tcaacatcac	caagacacac	atccaccagc	tcaccatgca	gaaggaagag	gtatccaggt	4500
gctacccggg	tcagatgata	aatctcacc	gagaaaaacg	gagcttatct	gaagaataaa	4560
agaggctgaa	gaacactctc	accagaccca	cagaaatctg	cagggaagtg	aaactgaaag	4620
tccaacagca	aaaggccact	ggctctgagg	tgtctcagag	gaaacagcag	ctcgaggttg	4680
agctgagaca	agtcacacag	atgcgaacag	aggagagcgt	aagatatga	caatctcttg	4740
atgatgctgc	caaaaccatc	caggataaaa	acaaggagat	agaaaggtta	aaacaaactga	4800
tcgacaagaa	aacaaatgac	cggaaaatgc	tggaagatga	aaacgcgaga	ttacaaaggg	4860
tccagatgta	cctgcagaaa	gcaaacagta	gtgcgcagga	gacaataaac	aaactgaaag	4920
ttcaggagca	agaactgaca	cgctcgagga	tgactatga	aagggtttcc	caggagagga	4980
ctgtgaagga	ccaggtatct	acgcggttcc	gaaaagctct	gaaaagctct	cgctgcagaa	5040
agcagaaggt	ggaagaggag	ctgaatcggc	tgaagaggac	cgctgcagaa	gactcctgca	5100
agaggaagaa	gctggaggaa	gagctggaag	gcagtgaggag	gtcgtgtaag	gagcaagcca	5160
tcaaaatcac	caacctgacc	cagcagctgg	acaggcgatc	catgtgtta	aagagggagt	5220
aggatgacct	ccggcagcag	agggacgtgc	tggtggcca	cctgagggaa	aagcagagga	5280

cccaggaaga	gctgaggagg	ctctctctct	aggctgaggc	cctgaggcgg	cagttactcc	5340
aggaacacga	aagtgtcaaa	caagctcact	tgaggaaatg	gcatttccag	aaggcgatag	5400
aaagataaa	cagaagctta	aatgaaagca	aaatagaaat	tgaggagctg	cagctctctca	5460
cagagaacct	gaccaaggtg	cacttgatgt	tagaagaaga	actcggggaac	ctgaggtctgg	5520
agtcacgatg	ctcgaggaga	ggacgaagcg	aagcggacag	tgataaaaaat	gcaacctctct	5580
tggaactaag	gagccagctg	cagatcagca	acaaccggac	cctggaactcg	cagggctctga	5640
ttaactgatt	acagagagag	agggaaaatt	tgagcacagga	aatttgagaaa	ttccaaaagc	5700
aggcttttaga	ggcattctaag	aggattcagg	aatcaagaag	tcagttgact	caggtgttac	5760
aggaagagaga	gagccttctg	gtgaaaatca	aagtcctcga	gcaagacaag	gcaaggtctgc	5820
agaggctgga	ggatgagctg	aatcgtgcaa	aatcaactct	agaggcgagaa	accaggtgta	5880
aaacagcgctc	ggagctgtgag	aaacagcaaa	ttcagaatga	cctgaatcag	tggaagactc	5940
aatattcccg	caaggaggag	gctattagga	agatagaatc	ggaagagagaa	aagagtgaga	6000
gagagaagaa	cagctcttag	agtgagatcg	aaagactcca	agcagagatc	aagagaattg	6060
aagagaggtg	caggcgttaag	ctggaggatt	ctaccaggga	gacacagtc	cagttagaaa	6120
cagaacgctc	cogatatcag	agggagattg	ataaactcag	acagcgccca	ttgggtgtcc	6180
atcgagagac	gcagactcag	tgtgagtga	cogttagac	ctccaaagct	gtgttttgatg	6240
ggctgaggaa	gaaggtgaca	gcaatgcagc	tctatgagt	tcagctgctc	gcacaaaacaa	6300
ccttgagcaa	actatttgag	gggaagaagt	cagtggaaga	agttgtctct	gaaatccagc	6360
cattctctga	gggtgcagga	tctatcgctg	gagcatctcg	ttctctctaa	gaaaaactct	6420
ctttgtaga	ggcacaagga	aaagaattaa	tcagcccaga	ctccaaagct	atgcttctag	6480
agggccaggc	agctacaggt	ggtataattg	atcccatc	gaatgagaag	ctgactgtcg	6540
acagtgcctat	agctcgggac	ctcattgact	togatgacg	tcagcagata	ttctcagcag	6600
aaaaagctat	cactggtttt	gatgatccat	tttcaggcaa	gacagtatct	gtttcagagc	6660
ccatcaagaa	aaatttgatt	gatagagaaa	cogggaatcg	cctgctggaa	gcccagattg	6720
cttcagggag	tgtagttagc	cctgtgaaca	gtgtcttttt	gcacaaagat	gtcgctctgg	6780
cccgggggct	gattgataga	gatttgtatc	gatccctgaa	tgatccccga	gatagtcaag	6840
aaaaacttgg	ggatccagtc	acccaaaaga	aggctcagta	cgtgcagctg	agaagaaagct	6900
gcagaaatcga	accacatact	ggtctgctct	tgcttcaagt	acagaaagag	agcagtgtct	6960
ttccagggaat	cagacaacct	gtgaccgtca	ctgagctagt	agattctggt	atattgagac	7020
cgtccactgt	catgaagact	gaatctggtc	tgacagaggt	tgacagaggt	gttgtagagaa	7080
ttaaggactt	cctccagggt	ccaagctgca	tagcaggcat	atacaatgag	accacaaaac	7140
agaagcttgg	catttatag	gcoatgaaaa	ttggttagt	ccagactggt	actgctctgg	7200
agttgtctga	agoccaaagc	gctactggct	ttatagtgg	ttctgttagc	aacttctagg	7260
taccagtgga	ggaagcctac	aagagaggtc	tggtgggcat	tgagttcaaa	gagaagctcc	7320
tgtctgacga	acgagctgtc	actgggtata	atgatcctga	aacagggaac	atactctctt	7380
gtttccaagt	catgaataag	gaactcatcg	aaaaggccca	cggtagctgc	ttattagaag	7440
ccacagctgc	aaccgggggg	atcattgacc	caaaggagag	ccatcgttta	ccagttgaca	7500
tagcatataa	gaggggctat	ttcaatgagg	aactcagtg	gattctctca	gatccaagt	7560
atgataccaa	aggaattttt	gaocccaaca	ctgaagaaaa	tcttacctat	ctgcaactaa	7620
aagaagaagt	cattaagagc	gaggaaacag	ggctctgtct	ttgcctctcg	aaagaataaa	7680
agaaacaggt	gcagacatca	caaaagaata	ccctcaggaa	cgctagagtgt	gtcatagtgt	7740
accacagaag	caataaagaa	atgtctgttc	aggaggccta	caagaagggc	ctaatgtatt	7800
atgaaaacott	caaaagactg	tgtgagcagg	aatgtgaatg	ggaagaaata	accatcaocg	7860
gatcagatgg	ctccacacgg	gtggtctcgg	tagatagaaa	gacaggcagc	cagtatgata	7920
ttcaagatgc	tattgacaag	ggccttgttg	ggcagaagtt	ctttgacag	ttacgatccg	7980
cgagctcag	cctcactcaa	tttgctgaca	tgtactcctt	gaaaaatggt	gtcggccacca	8040
gcagcagcat	gggcagtggt	ctcagcgatg	atgtttttag	cagctccocga	catgataagc	8100
taagttaagt	ttccaccata	tcocagctca	ggaatttaac	cataaggagc	agctcttttt	8160
cagacaacct	gaaagaatcg	agccccattg	cagccatctt	tgacacagaa	aacctggaga	8220
aaatctccat	tacagaaggt	atagagcggg	gcactgttga	cagcatcacg	ggctcagagg	8280
ttctggaggc	tcaggcctcg	acaggtggca	tcactcaccc	aaccacgggc	cagaagctgt	8340
cacttcagga	cgcagctctc	cagggtgtga	ttgacacaga	catggccacc	agcgtggaag	8400
ctgtcgcagaa	agccttctca	ggcttcgagg	gtgtgaaggg	aaagaagaag	atgtcagcag	8460
cagagcgagt	gaaagaaaaa	ttgctccogt	atgagctgg	ccagcgttcc	ctgaggtctc	8520
agtaacctcac	gggaggtctt	gttgaccocg	aagtgcattg	gaggataaag	accgaagaag	8580
ccatccggaa	ggggttctca	gatggcccg	ccgcacagag	gctgcaagac	accagcagct	8640
atgccaaaat	cctgactcgc	cccaaaaaca	aattaaaaat	atctctataag	tgatgcataa	8700
atcgctccat	ggtagaagat	atactggg	tgccgtctct	ggaagccgcc	tcogtctcgt	8760

```

ccaagggtctt acccagccct tacaacatgt cttcggtctc ggggtccgcg tccggtctcc 8820
gtctgggtatc tgcgtccgga tctcgtccg ggtccgcag tgggtccgcg agaggaagct 8880
ttgaagccac aggggaattct tctactctt attcctactc atttagcagt agttctattg 8940
ggcactag                                     8948

```

```

<210> 120
<211> 587
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 91, 131, 256, 263, 332, 392, 400, 403, 461, 496, 497, 499,
510, 511, 518, 519, 539, 554, 560, 576
<223> n = A,T,C or G

```

```

<400> 120
ggtcctaagc acttagacta catcaggga gaacacagac cacatccctg tctcatgctg 60
gottatgttt tctggaagaa agtggagacc nagtccttgg ctttagggct ccccggtctg 120
gggctgtgca ntccggtcag ggcgggaagg gaaatgcacc gctgcatgtg aacttacagc 180
ccaggcggaat gccctctccc ttagcactac ctggcctcct gcacccctc gctcatgtt 240
cctcccacct tcaanaaatg aanaaccoca tgggccacgc ccttgccctt ggggaaccaa 300
ggcagccttc caaaactcag gggctgaagc anactattag ggcaggggct gactttgggt 360
gacactgccc attcctctct agggcagctc angtcacccn ggnetcttga acccagcctg 420
ttcctttgaa aaagggcaca actgaaaagg gcttttctta naaaagaaa aaccaggga 480
ctttgcaggg gcttcnntnt taccaaaacn ncttctcnn gatttttaat tcccattng 540
gctccactt accnggggcn atgccccaaa attaanaatt tcccatc 587

```

```

<210> 121
<211> 619
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 260, 527, 560, 564, 566, 585, 599
<223> n = A,T,C or G

```

```

<400> 121
cactagtagg atagaaacac tgtgtccoga gagtaaggag agaagctact attgattaga 60
gcctaaccca ggttaactgc aagaagagcg gggatacttt cagctttcca tgtaactgta 120
tgcataaagc caatgtagtc cagtttctaa gatcatgttc caagtaactt gaatccact 180
tcaatacaca ctcatgaact cctgatggaa caataacagg cccaagcctg tggtagtgatg 240
tgacactttg ctgagctcan aaaaaatact actctcataa atgggtggga gtattttgg 300
gacaacctac tttgcttggc tgagtgaagg aatgatattc atatattcat ttattccatg 360
gacatttagt tagtgctttt tatataccag goctgatgct gagtgaactt cttgtgtata 420
tttccaaatt tttgtacagt cgtgcacat attgaaatc atatattagc acttccaaaa 480
aatgaagctc ctggtttttc atggcaactt gatcagtaaa ggattcncct ctgtttggta 540
cttaaaacat ctactatnct gtttnanatg aattcctttt cccnccctcc cgaaaaaana 600
aagtgtggg gaaaaaaa                                     619

```

```

<210> 122
<211> 1475
<212> DNA
<213> Homo sapiens

```

```

<400> 122
tccacctgtc cccgcagcgc cggctcgccg cctcctgcgc cagccaccga gccgccgtct 60
agcgcccccga cctgcgccacc atgagagccc tgcctggcgcg cctgcttcctc tgcgtcctgg 120
tgcgtgagcga ctccaaagcgc agcaatgaac ttcactcaagt tccatogaaac tgtgactgtc 180
taaatggaggg aacatgtgtg tccaaacaagt acttctccaa cattcactcgg tgcaactgcc 240
caaaagaattt cggaggggcag cactgtgaaa tagataagtc aaaaacctgc tatgagggga 300
atgggtcactt ttaccggagga aaggccagca ctgacaccat gggccggccc tgccctgccct 360
ggaaactcgc cactgtccct cagcaaaogt accatgccca cagatctgat gctcttcagg 420
tggggcctggg gaaacataat tactgcagga acccagacaa ccggaggcgga cctgtgtgct 480
atgtgcaggtg gggcctaagc ccgctgtgtcc aagagtgcatt ggtgcattgac tgcgcagatg 540
gaaaaaagcc ctccctctcct ccagaagaat taaaatttca gtgtggccaa aagactctga 600
ggccccgctt taagattatt gggggagaaat taccacacat cgagaaccag ccctggtttg 660
cggccactcta caggagggcac cggggggggct ctgtcaccta cgtgtgtgga ggcagccctca 720
tcagcccttg ctgggtgatc agcgccacac actgottcat tgattaccca aagaaggagg 780
actacatcgt ctacctgggt cgctcaaggc ttaactccaa cagcaagggg gagatgaagt 840
ttgaggttga aaacctctat ctacacaagg actacagcgc tgacacgctt gctcaccaca 900
acgacattgc cttgctgaag atccgttcca aggaggggcag gtgtgcgcag ccatcccgga 960
ctatcacgac cactctgcct cctcgtatgt taaacgatcc cagttttggc acaagctgtg 1020
agatcactgg ctttggaaaa gagaatttcta ccgactatct ctatccggag cagctgaaga 1080
tgactgttgt gaagctgatt tcccaccggg agtgtcacga gcccactac tacggctctg 1140
aagtcaccac caaaattgctg tgtgtgctgt accccacgtg gaaaacagat tctcgcagg 1200
gagactcagg gggaccctct gctgttcccc tccaaaggcgc catgactttg actggaattt 1260
ctgactgggg gccctgaagt gccctgaagg ctaagccagg cgtctacacg agagtctcac 1320
actcttacc ctggatccgc agtcacacca aggaagagaa tggcctggcc ctctgagggt 1380
cccacggag gaaaacggca ccaccgctt tctgtcgtgt tgtcattttt gcagtagagt 1440
catctccatc agctgtaaga agagactggg aagat 1475

```

```

<210> 123
<211> 2294
<212> DNA
<213> Homo sapiens

```

```

<400> 123
cagcgccggc tgcgcgcctc ctgcgcgagc caccgagcgc ccgtctagcg ccccgacctc 60
gccaccatga gagcctcgtc ggcgcgccgt cttctctgcg tctgtgtcgt gagcgactcc 120
aaaggcgaca atgaacttcca tcaagttcca tcgaactgtg actgtctaaa tggaggaaaca 180
tgtgttcacca acaagtaact ctccaacatt cgtcgtgtgca actgcctgca agtgcgccaa gaatttcgga 240
gggcagcact gtgaaataga taagtcaaaa acctgctatg aggggaatgg tcaactttac 300
cgaggaaagg ccagcactga caccatgggc cggccctgcc tgccctggaa ctctgcaact 360
gtccttcagg aaacgtacca tgcccacaga tctgatgtgc tctcagtcgt cctcgggaaa 420
cataattact gcaggaaacc agacaaccgg aggcgacccct ggtgctatgt gcagggtggc 480
ctaaagccgc ttgtccaaga gtgcattggt catgactgcg cagatggaaa aagtcctcc 540
tctcctccag aagaattaaa atttcagtgt ggcacaaaga ctctgaggcc ccgttttaag 600
attattgggg gagaattcac caccatcgag aaccagccct ggtttgcggc catctacagg 660
aggcaccggg ggggctctgt caccatcggt tgtggaggga gcctcatcag ccttgcgtg 720
gtgatcagcg ccacacactg ctctcattgat taccacaaaga agggagcata catcgctac 780
ctgggtcgct caaggctttaa ctccaacacg caaggggaga tgaagtttga caccacaaa 840
ctaactctac acaaggacta cagcgctgac acgcttgctc accacaacga cactgactat cagtccttg 900
ctgaagatgc gttccaagga ggccaggtgt gcgcagccat ccoggactat accgaccat 960
tgctcgccct cgatgtataa cgatccccc atttggccaa gctgtgagat cactggcttt 1020
ggaaaagaga attctacaga ctatctctat ccggagcagc cggaaatgac tgttgtgaag 1080
ctgatttccc accgggagtg tcagcagccc cactactcag gctctgaagt caccacaaa 1140
atgctgtgtg ctgctgaccc acagtggaaa acagattcct gccagggaga ctcaggggga 1200
ccctcgtctc gttccctcca aggcgcgatg actttgactg gaattgtgag ctggggccgt 1260
ggatgtgccc tgaaggacaa gccaggcgctc tacacgagag tctcacactt cttacctgg 1320
atccgcagtc acaccaagga agagaatggc ctggccctct gaggggtccc agggaggaaa 1380
cggggcaccac ccgctttctt gctggttgct attttcaggt agactcatct cactcagctg 1440
taagaagagc tgggaatata ggctctgcac agatggattt gcctgtgcca ccaccagggc 1500

```

gaacgacaat	agctttaccc	toaggcatag	gcctgggtgc	tggtgcccac	gacccctctg	1560
gccaggatgg	agggggtggt	ctgactcaac	atgttactga	ccagcaactt	gtctttttct	1620
ggactgaagc	ctgcaggagt	taaaaagggc	agggcatctc	ctgtgcatgg	gctcgaaggg	1680
agagccagct	ccccccagcg	gtgggcattt	gtgaggccca	tggttgagaa	atgaataatt	1740
tcocaaattag	gaagtgttaag	cagctgaggt	ctcttgaggg	agcttagcca	atgtggggagc	1800
agcggttttg	ggagcagaga	cactaacgac	ttcaggggcag	ggctctgata	ttccatgaat	1860
gtatcaggaa	atatatatgt	gtgtgtatgt	ttgcacactt	gtgtgtgggc	tgtagtgta	1920
agtggtgagta	agagctgggt	tctgtattgtt	aagtctaaat	atttcottaa	actgtgtgga	1980
ctgtgatgcc	acacagagtg	gtctttctgt	agaggttata	ggctactctc	ggggcctctt	2040
gggtccccc	cgtgacagtg	cctgggaatg	tattattctg	cagcatgacc	tgtgaccacg	2100
actgtctcag	tttcaacttc	acatagatgt	ccctttcttg	gccagtttat	cottcctctt	2160
agcctagttc	atccaatcct	cactgggtgg	ggtgaggacc	actcctgtac	actgaattatt	2220
tatatattcac	tatttttatt	tatatatttg	taatttttaa	taaaagtgat	caataaaatg	2280
tgatttttct	gatg					2294

&lt;210&gt; 124

&lt;211&gt; 956

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 124

gatgagttcc	gcaccaagtt	tgagacagac	caggccctgc	gcctgagtgt	ggaggccgac	60
atcaatggcc	tgcgcaaggt	gctggatgag	ctgaccctgg	ccagagccga	cctggagatg	120
cagattgaga	acctcaagga	ggagctggcc	tacctgaaga	agaaccaaga	ggaggagatg	180
aacgccctgc	gaggccaaggt	gggtggtgag	atcaatgtgg	agatggacgc	tgccccaggc	240
gtggacctga	gcgccatcct	caacgagatg	cgtgaccagt	atgagaagat	ggcagagaa	300
aaccgcgaag	atgocgagga	ttggtctctc	agcaagacag	aggaactgaa	cgcgaggggt	360
gccaccaaca	gtgagctggt	gcagagtggc	aagagtgaga	tctcgagctc	cggcgccacc	420
atgcaggcct	tggagataga	gctgcagctc	cagctcagca	tgaagcagtc	cctggagggc	480
aacctgtcgc	agacagagaa	cgcctactgc	gtcagcgtgt	ccagatccca	ggggctgatt	540
ggcagcgtgg	aggagcagct	ggcccagctt	cgtcgcgaga	tgagagcgca	gaaccaggaa	600
tacaaaatoc	tgctggatgt	gaagacgcgg	ctggagcagg	agattgccac	ctaccgcgcg	660
ctgctggagg	gcaggagtgc	ccacctgact	cagtaacaaga	aagaaccggt	gaccaaccgt	720
cagggtcgta	ccatttgtga	agaggtccag	gatggcaagg	tcactctctc	cgcgagcgag	780
gtccaccaga	ccaccgcgtg	aggactcagc	taccocggcc	ggccaccagc	gaggcagcca	840
cgcagccgcg	ccatctgccc	cacagtctcc	ggcctctcca	gcctcagccc	cctgcttcag	900
tcctctcccc	atgcttctct	gcctgatgac	aataaaagct	tggtgactca	gctatg	956

&lt;210&gt; 125

&lt;211&gt; 486

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 16

&lt;223&gt; n = A, T, C or G

&lt;400&gt; 125

aaattatata	tagtgnttca	gtcccccattg	tggtgtttcat	agtcttctag	gaacagataa	60
acttaagtat	tcaattcaat	cttggccattt	ttctctttaa	ataggctttt	tagcctattt	120
ttggaaaaat	gctttttctc	tgagaacctt	ttcttgaatg	tcatacaatt	taccaaaacct	180
tctaatgtoca	gagctaaatt	agtaactgtt	aagttactat	tgactgaatt	ttcttcaatt	240
tctgttttagc	cagtgtttacc	aaagtaagct	ggggaatgaa	gtataccaac	ttctttcaga	300
gcatttttagg	acattatgac	agcttttaga	ggctgtcttg	ttcttagcca	agggagagcc	360
agcgcgaggt	ttggataacta	gagaaagtca	tttgcttgta	ctattgccat	tttagaaagc	420
tctgatgtga	attcaaaatt	taacctctgt	acttaaagcc	aacaatttta	aggcagtagt	480
ttttat						486

<210> 126  
 <211> 3552  
 <212> DNA  
 <213> Homo sapiens

<400> 126

```

cgccaggcag  gctcgtctc  ggcaccctcc  cgccgccccc  gttctcctgg  cctcggccgg  60
catcccgatg  gccgccgctg  ggcgccggcg  ctcgtgctgc  ggagcgcgtc  gccctcatct  120
gctgctgacc  ctctgtgatc  tcaagtctgc  tgggtaagcg  tcgcaaaaagg  tgatacttaa  180
tgtactctct  aaactagagg  cagacaaaaa  aattggcagä  gtaattttgg  aagagtgtct  240
caggctctgca  gacctcatcc  ggtaaatgga  tctcgtattc  agagtcttaa  atgatgggtc  300
agtgtacaca  gccaggcgct  ttgcgctgto  tgataagaaa  agatcattta  ccatatggct  360
ttctgacaaa  aggaacacga  cacagaaaga  ggttactgtg  ctgctagaac  atcagaagaa  420
ggtatcgagg  acaagacaca  ctagagaaac  tgtttctcagg  cgtgccaaaga  ggagatgggc  480
acctattcct  tgcctatagc  aagagaatto  cttgggcctc  ttccattgt  ttctccaaca  540
agttgaatct  gatgcagcac  agaactatac  tgtcttctac  tcaataaagt  gacgtggagt  600
tgataagaaa  cctttaaatt  tgttttatat  agaaagagac  actggaaatc  tatittgcac  660
tcggcctgtg  gatcgtgaag  aatatgatgt  ttttgatttg  attgcttatg  cgtcaaatgc  720
agatggatct  tccagcagat  tgcccctccc  actaccctac  agggtagagg  atgaaaaatga  780
caaccacctc  gttttcacag  aagcaattta  taattttgaa  gttttggaaa  gtatgagacc  840
tggtaactaca  gtgggggtgg  tttgtgccac  agacagagat  gaaccgggac  caatgcatac  900
ggcgtgaaa  tacagcattt  tgcagcagac  accaaggctc  cctcgtctct  tatittgcac  960
tcccagaca  ggcgtaatac  ccacagcttc  tcaattattg  gacagagagg  ttgtagacaa  1020
gtactactgc  ataatagaag  tacaagacat  ggaatggccag  ttttttgat  tgataggcac  1080
atcaacttgt  atcataacag  taacagattc  aaatgataat  gacccactt  tcagacaaaa  1140
tgcttatgaa  gcatttgtag  agaaaaatgc  attcaattgt  gaaatcttac  gaataacctat  1200
agaagataag  gatttaatta  acactgocaa  ttgggagctc  aattttacaa  ttttaaaggg  1260
aaatgaaatt  ggacatttca  aaatcagcac  agacaaagaa  actaatgaag  gtgtctcttc  1320
tgtttaaag  cactgtaatt  atgaagaaaa  ccgtcaagtg  aaactggaaa  ttggagtaaa  1380
caatgaagcg  ccaatttgct  gagatattcc  cagagtgaca  gccttgaaca  gaggccttgt  1440
tacagttcat  gtgagggatc  tggatgagg  gctgtaatgc  actcctgcag  cccaatatgt  1500
goggattaaa  gaaaacttag  cagtggggtc  aaagatcaac  ggctataagg  catatgacc  1560
cgaaaaataga  aatggcaatg  gtttaaggta  caaaaaatg  catgatccta  aaggttggtat  1620
caccattgat  gaaatttccg  ggtcaatcat  aacttccaaa  atcctggata  gggaggttga  1680
aactcccaaa  aatgagttgt  ataattattc  agctcgtgca  atagacaaag  atgatagatt  1740
atgtactgga  acactgtgct  tgaacattga  agatgtaaa  gataatccac  cagaaatact  1800
tcaagaatat  gtatgcattt  gcaaacccaa  aatgggggat  accgacattt  tagctgttga  1860
tctcgtatgaa  cctgtccatt  gagctccatt  ttatttcagt  ttgcccaata  ctctccaga  1920
aatcagtaga  ctgtggagcc  tcaccaaaat  taatgataca  gctgcgcgtc  ttccatatac  1980
gaaaaatgct  ggattttcaa  aataatccat  tctattact  gtaaaagaca  gggcgccgca  2040
agctgcacaa  aaattatgga  gagttaatct  gtgtgaatgt  actcatocaa  gttccattgc  2100
tgcgacttca  aggagtacag  gagtaatact  tggaaaatgg  gcaatccttg  caatattact  2160
gggtatgcaa  ctgctctttt  ctgtattgct  aacttttagt  tgtggagttt  ttgttgcaac  2220
taaaaggaaa  gctttctcgt  aagatttagc  acagcaaaac  ttaattatat  caaacacaga  2280
agcacctgca  gacgatagag  tgtgctctgc  caatggattt  atgacccaaa  ctaccaacaa  2340
ctctagocaa  ggtttttgtg  gtactatggg  atcaggaatg  aaataatgca  gctcgtgact  2400
cattgaaatg  atgaaggagg  gaaaccagac  cttggaatcc  tgccgggggg  ctgggcatac  2460
tcataccctg  gactcctgca  ggggaggaca  caccaggtg  gaccaactga  gatacactta  2520
ctcggagtgg  cacagtttta  ctcaaccccg  tctcgtgtaa  aaattgcact  gaattgaatc  2580
gaatgaagac  cgcgatccat  cccaagatta  tgtctcact  tataactat  agggaaagat  2640
atctccagct  gctcgtcgag  gctgctcgag  tgaaaagcag  gaagaagatg  gctcgtgact  2700
tttaaatat  ttggaaccca  aattttattc  attagcagaa  gcatgccaaa  agagataatt  2760
tcacagtgt  acaattaggt  ctttgcaga  cattctggag  gtttccaaaa  ataatttgt  2820
aaagtccat  ttcaacatgt  atgtatatga  tgattttttt  ctcaattttg  aattatgcta  2880
ctccaacatt  tatattttta  aagcaagttg  ttgcttatct  ttcccaaaaa  gtgaaaaatt  2940
ttaaaaacga  caactggtaa  atctcaacct  ccagcactgg  aattaaggtc  tctaaagcat  3000
ctgctctttt  ttttttttac  agatatttta  gtaataataa  tgctggataa  atatttagtc  3060

```

```

aacaatagct aagttatgct aatatcacat tattatgtat tcactttaag tgaatgttta 3120
aaaaataaac aagaaatatt gagtatcact atgtgaagaa agttttggaa aagaacaat 3180
gaagactgaa ttaaatataa aatgtttgcag ctcataaaga attggactca ccctactgc 3240
actaccaaatt tcattttgact ttggagggcaa aatgtgttga agtgccctca gaagtgcgaa 3300
ttttctatag gaatatagtt ggaaataaat gtgtgtgtgt atattattat taatcaatgc 3360
aatatttataa tgaaatgaga acaaaagaga aaatggtaaa aacttgaat gaggtctggg 3420
tatagtttgt cctacaatag aaaaaagaga gagcttccca ggctgtggct cttaaatgct 3480
gcattataac tgagtctatg aggaaatagt tccgtcccaa ttgtgtgaat ttgtttaaaa 3540
ttgtaataaa at                                     3552

```

<210> 127  
 <211> 754  
 <212> DNA  
 <213> Homo sapiens

```

<400> 127
tttttttttt ttgtcattgt tcattgattt taatgagaaa gctaagagag gaaataagta 60
gcctttcaaa ggtcacacag aagtaagtga cagatccagg attccatctc aagcattctg 120
gctctagtgt ccatgcttct caaccattat gaccacaat tcaaccacaat caatactgaa 180
ggacaagtga aatgtatccg gtatttttact attacaaaca aaaatccaat gaacattctt 240
gaagacatac acaaaaaataa tgggttacaat agaagttact ggaattgaaa ttttgggtca 300
acctatatta aaatgtgaag cttttgatat agctaataga tttttgaaat gatcagttct 360
aacgtttgta ggggagcaca ctccgtgatg gggaaaaagat tccactgtgaa gcacagagca 420
cctttatggt tggatcatct tgtcattaaa gtccagcgct tatctatcct gtaagtggca 480
gaatcaagac tgcaatatcg cctgcttttc tttttaactc atgttttccc ttgactacac 540
tgggtcctcaa agtaaaaccc ctgtgtcagt gtaactattca tggaaatactc tgcattata 600
accaccttct aatactttta ataccacatc aaaatttatt atacatatgt atcatagata 660
ctcatctgta aagctgtgct tcaaaatagt gatctcttcc caacattaca atatatatta 720
atgatgtcga acctgccccg gcggcccgctc gaag                                     754

```

<210> 128  
 <211> 374  
 <212> DNA  
 <213> Homo sapiens

```

<400> 128
aggttttgat taaaaaggca aatgatttta ttgttcgata atctttttaa aaaataagag 60
gaaggagtaa aattaaagat gaaagatgat ttttatctcc ttgtgacctc tatatccccc 120
ttcccttgcc cttggtaagt aactcctgat ggagaaagga ttaagaacctc ttatttaacc 180
aaaaaacaga gccagctaat catttccaaa ggttagtatc tccctgtgga cctctctctt 240
ggtttaattg aataaaacta tatgttcata tatgtattaa acaactcag aataacatct 300
ttctctcctt agttaaggca ttataagggc tatactatca tccataataa ccaaggccaat 360
aacctaaaaa gctg                                     374

```

<210> 129  
 <211> 546  
 <212> DNA  
 <213> Homo sapiens

```

<400> 129
agtgtgatgg atactctgag aattcgggct aagcgtggtc gcggcccgag gtctggaact 60
tcccagcacy tgaaaaaggag cctcctgagc tgactcggct aaagcccccac tttcgtctct 120
cctcatttct gccctactgat ttccctggag cattcactcg aatattaccg tttgtctgtg 180
aacctgtgac atacatagca tgactccctg gaatagagtg ggcctgggtg cttatgctgg 240
gagagtgtat gacatgcact ttcaagctat atccaccatt tgcagcaaaag gaaaaaaaat 300
acctcgagta aattccatca ttttttataa catcagcacc tgctccatca tcaaggagtc 360
tcagcgttaac aggatctcca gtctctggct caactgtggt agtgacagtg gcatcaagaa 420
tgggataaaa tccctgtttc acattggcat aaatcatcac aggatgagga aaatggaggc 480

```

tgtctctttt cacaagggt tccacagtgg ctgggggcac agacctgccc gggcgccgc 540  
 tcgaaa 546

<210> 130  
 <211> 5156  
 <212> DNA  
 <213> Homo sapiens

<400> 130  
 accaacggag ggcggggca ggcacccctg cagcggagac agagactgag cggcccgcca 60  
 ccgcatcggt tgcgctctgg ctggggtgct gacctgtgct gtgcgtctcc ctgcccgcag 120  
 ccggggccac ctccaggagg gaagtctgtg attgcaatgg gaagtcagg cagtgtatct 180  
 ttgatcgagg acttccacaga caaactggta atggattccg ctgcctcaac tgcattgaca 240  
 acactgatgg cattcaactgc gagaagtcca agaattggct ttaccggcac agagaaagg 300  
 accgtctgtt gccctgcaat tgaactcca aaggttctct tagtgctcga tgtgacaact 360  
 ccggacgggt cagctgtaaa ccagggtgta caggagccag atcgaccca tgtctgcag 420  
 gcttccacat gctcacggat gcgggggtgca ccaagacca gacagtcta gactccaagt 480  
 gtgactgtga cccagctggc atcgaggggc cctgtgacgc gggcgctgt gtctgcaagc 540  
 cagctgtcac tggagaacgc tgtgataggt gtcgatcagg ttactataat ctggatgggg 600  
 ggaacctgga gggctgtacc cagtgtttct gctatggcca ttcagccagc tggccgagct 660  
 ctgcagaata cagtgctcat aagatcacct ctaccttca tcaagatgtt gatggctgga 720  
 aggctgtcca acgaaattgg tctcctgcaa agctccaatg gtcacagcgc catcaagatg 780  
 tgtttagctc agcccaacga ctgacacctg tctattttgt ggctcctgoc aaatttcttg 840  
 ggaatcaaca ggtgatctat ggtcaagcc tgcctttgta ctaccgtgtg cacagaggag 900  
 cgacacccc atctgcccac gatgtgattc tggaaaggct tggctctacgg atcacagctc 960  
 ccttgatgcc acttggcaag acactgcctt gtgggctcac caagacttac acattcaggt 1020  
 taaatgagca tccaaagcaat aattggagcc ccagctgag ttactttgag tatcgaaagt 1080  
 tactcgagaa tctcacagcc ctccgcatca gacatacata tggagaatac agtactgggt 1140  
 acattgacaa tgtgacctgt atttcagccc gccctgtctc tggagcccca gcaacctggg 1200  
 ttgaacagt tatatgtcct gttgggtaca aggggcaatt ctgccagat tgtgttctg 1260  
 gctacaagag agattcagcg agactggggc cttttggcac ctgtattcct tgaactgtc 1320  
 aagggggagg ggctgtgat ccagacacag gagattgta ttcaggggat gagaatcctg 1380  
 acattagtt tgcctagctc ccaattgggt tctacaacga tccgacagac ccccgagct 1440  
 gcaagccatg toctgtcat aacgggttca gctgctcagt gatgcggag acggaggagg 1500  
 tgggtgcaa taaactccct ccgggggtca ccggtgcccg ctgtgagctc tgtgtgatg 1560  
 gctaacttgg ggaccccttt ggtgaacatg gccagtgag gcctgtgcag cctgtcfaat 1620  
 gcaacaacaa tgtgacccc agtgctctg gaaattgta ccggctgaca cggcagtgatt 1680  
 tgaagtgtat ccacaacaca gccggcatct actgcgacca gtgcaagca ggctacttgc 1740  
 gggaccatt ggctcccaac ccagcagaca agtgtcgagc ttgcaactgt aacctatgt 1800  
 gctcagaggt tgtaggatgt cgaagtgtg gcacctgtgt ttgcaagcca gattttgggt 1860  
 gccccaact tgaagcatga gcaatcagct gtcagctgt ctataatcaa gtgaagattc 1920  
 agatggaatc gtttatcgag cagcttcaga gaatggaggc cctgatttca aagctcagg 1980  
 gtggtgatgg agtagtaact gatacagagc tggaaaggag gatgcagcag gctgagcagg 2040  
 cccctcagga cattctgaga gatgcccaga ttccagaagg tgctagcaga tcccttggtc 2100  
 tccagttgag caaggttgagg agccaaagga gagcgtacca gagccgctg gatgaactca 2160  
 agatgactgt ggaagaagtt cgggctctgg gaagtcaatg ccagaaccca gcttcgggata 2220  
 ctccagcgt catcactcat atgcagctga cctggcaga aagtgcagca gctttgggaa 2280  
 acactaaact tccctgcctca gaccactacg tggggccaaa tggctttaaa agtctggctc 2340  
 agggggccac aagattagca gaaagccaag ttgagtcagc cagtaaacatg gagcaactga 2400  
 caagggaaac tgaggactat tccaaacaag cccctcact ggtgogcaag gccctgcatg 2460  
 aaggagtcgg aagcgaagcg ggtagcccg acggtgtgct ggtgcagagg cttgtggaaa 2520  
 aattggagaa aaccaagctc ctggcccagc agttgacaag ggaggccaat caagcgga 2580  
 tgaagcaga taggtcttat cagcacagtc tccgctcct gattcagtg tctgcgtctc 2640  
 agggagtcag tgatcagctc tttcaggtgg aagaagcaaa gaggatcaaa caaaagcggc 2700  
 attcaactct aagcctggtt accaggcata tggatgagtt caagcgtaca cagaagaatc 2760  
 tgggaaactg gaaagaaaga gcacacagc tcttacagaa tggaaaaagt gggagagaga 2820  
 aatcagctca gctgtcttcc cgtgccaact ttgttaaaag cagagcaaaa gaggcaactg 2880  
 gtatgggcaa tgcactttt tatgaagttg agagcatcct taaaaacctc agagagtttg 2940



acctgcagg	ggacaacaga	aaagcagaag	ctgaagaagc	catgaagaga	ctctcctaca	3000
tcagccagaa	ggtttcagat	gccagtgaca	agaccagaca	agcagaaaga	gccttgggga	3060
gcgctgctgc	tgatgcacag	agggcaaaaga	atggggcccg	ggaggccctg	gaaattctcca	3120
gtgagattga	acaggagatt	gggaggtctga	acttggaaagc	caatgtgaca	gcagatggag	3180
cccttggccat	ggaaaaaggga	ctggccctctc	tgaagagtgga	gatgaggggaa	gtggaaggag	3240
agctgggaaa	gaaggagctg	gagtttgaca	cgaatatgga	tgcatgacag	atgggtgatta	3300
cagaagccca	gaagggtgat	accagagcca	agaacgctgg	ggttacaatc	caagacaac	3360
tcaacacatt	agacggcctc	ctgcactctga	tggaccagcc	tctcagtgta	gatgaagg	3420
ggctgggtct	actggagcag	aagctttccc	gagccaagac	ccagatcaac	agcaactggc	3480
ggcccatgat	gtcagagctg	gaagagaggg	cacgtcagca	gaggggccac	ctccatttgc	3540
tggaagacaag	catagatggg	attctggctg	atgtgaagaa	cttggagaa	attagggaca	3600
acctgcccc	aggctgctac	aataccagag	ctcttgagca	acagtgaagc	tgccataaat	3660
atttctcaac	tgaggttctt	gggatacaga	tctcagggct	gggagcccat	gtcatgtgag	3720
tggtgggtgat	ggggacattt	gaacatgttt	aatgggtatg	ctcaggtcaa	ctgacctgac	3780
ccattctctg	atcccatggc	caggtgggtg	tcttatttga	ccatactcct	tgcttctctga	3840
tgctgggcaa	tgaaggcagat	agcactgggt	gtgagaatga	tcaaggatct	ggaccccaaa	3900
gaatagactg	gatggaaaga	caaaactgcac	aggcagatgt	ttgcctcata	atagtgttaa	3960
gtggagctct	ggcaatttga	caagtgctgt	tggtgatata	tcaacttatt	ctttgagtaa	4020
tgtgactaaa	ggaaaaaact	ttgactttgc	ccaggcatga	aattcttctc	aatgtcaaga	4080
cagagtgcaca	ccagtgccca	ctgtggccag	taaaatacta	ttgcctcata	ttgtcctctg	4140
caagcttctt	gctgatcaga	gttctctcta	cttacaaccc	aggggtgtgaa	catgtttctc	4200
attttcaagc	tggaagaagt	gagcagtggt	ggagtggaga	cctgtaaagg	agggccattc	4260
agagctatgt	tgcttctgtg	tgccctgccac	cttcaagttc	tggaactggg	catgacatcc	4320
tttcttttaa	tgatggcatg	gcaacttaga	gattgcattt	ttattaaagg	atttctacc	4380
agcaaaagcaa	atggttggaa	agtatttact	tttctgggtt	caaagtgaata	gaaaagtgtg	4440
gcttgggcat	tgaagaaggt	aaaattctct	agatttatta	gtcctaattc	aatcctaact	4500
ttagaacacc	aaaaatgatg	cgcatcaatg	tattttatct	tattttctca	atctcctctc	4560
tctttctctc	accataataa	agagaatgtt	ctcatcaaca	cttcagctgg	gctcaatcca	4620
tcctctccat	catccttcca	tccatcttct	catccattac	ctccatccat	ccttccaaca	4680
tatatattat	gagtaacctg	ttgtgtccag	gggtggttgg	gacagtggtg	acatagctctc	4740
tgccctcata	gagttgattg	tctagtggag	aagacaagca	tttttaaaaa	ataaatttaa	4800
acttacaac	tttgtttgtc	acaagtggtg	tttattgcga	taaccgcttg	gtttgcaacc	4860
tottttgtca	acagaacata	tggtgcaaga	ccctccatg	ggggcaactg	agttttggga	4920
aggctgacag	agctctgggt	tggtgcacatt	tctttgcatt	ccagctgtca	ctctgtgctc	4980
ttctacaact	gattgcaaca	gactgttgag	ttatgataac	accagtggtg	attgctggag	5040
gaaccagagg	cacttccacc	ttggctggga	agactatggt	gtgccttgc	ttctgtattt	5100
ccttggaatt	tcttgaaggt	gtttttaaat	aaagaacaat	tgtagaaaa	aaaaaa	5156

&lt;210&gt; 131

&lt;211&gt; 671

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 131

aggtctggag	ggccacacgc	cggatgtggg	acaccgggaa	aaagtgtgtca	tagcacacat	60
ttttgcatoc	cggttgcagt	gtgtgtcaga	cgaagctctc	ttgctcgtca	ccccacacct	120
cctgggcagc	caycacgagg	atcatgactc	ggaaaataaa	gatgactgtg	atccacacct	180
tcoccatgct	ggtggagtgt	ttgttgacac	ccoccatgaa	agtggtgcagc	gtcccccaat	240
ccattggcgt	gggttatccc	tgagtcctgt	ttccaacgac	tgccagtgtt	tcagacccaa	300
agaatgagct	caagatccct	ctgcgagggt	ttcagacctc	cttctcctac	ccactgtggag	360
tgcttagaag	ccaatgggtg	cacagtgatg	atacgaatgt	caattcttgc	tcggtcagtg	420
aggatgtcgc	ctggaatatt	caaattgaat	tacagatgca	tgaagagggc	tacaagtgtt	480
gaatttttct	ttcgccatac	agaaattgtt	tagccagatc	ttctgtactt	cttttctctc	540
cctgacacct	cctgtctccc	aggaagggag	gtcagccccc	tttgcaaaac	acaggatgcc	600
cgtgacaacg	gagacagggt	ttcttccacg	acaggaagtg	ccttctgggt	cctgcacgtt	660
ttactgcta	t					720

<210> 132  
 <211> 590  
 <212> DNA  
 <213> Homo sapiens

<400> 132  
 ctgaatggaa aagcttatgg ctctgtgatg atattagtga ccagcgggaga tgataagctt 60  
 ctgtgcaatt gcttaccacc tgtgtctcagc agtgggttcaa caattcactc cattgcccctg 120  
 ggttcaatctg cagcccccacaa tctggaggaa ttatcacgtc ttacaggagg tttaaagttc 180  
 tttgttccag atatatcaaa ctccaatagc atgattgatg ctttcagtag aatttccctc 240  
 ggaactggag acatttttoca gcaacatatt cagcttgaaa gtacagggtga aaatgtcaaa 300  
 cctcaccatc aattgaaaaa cacagtgact gtggataata ctgtgggcaa cgacactatg 360  
 tttctagtta cgtggcaggc cagtggctcct cctgagatta tattatttga tctctatgga 420  
 cgaaaaact acacaaataa ttttatcacc aatctaact ttoggacagc tagtcttttg 480  
 attccaggaa cagctaagcc tgggcaactg acttacacc tgaacaatac ccatcattct 540  
 ctgcaagccc tgaagtgac agtgacctct cgcgcctcca actcagacct 590

<210> 133  
 <211> 581  
 <212> DNA  
 <213> Homo sapiens

<400> 133  
 aggtcctgtc cgggggcact gagaactccc totggaattc ttgggggggtg ttggggagag 60  
 actgtggggcc tggagataaa acttgtctcc totaccacca cctgtgacc tagcctgac 120  
 ctgtctcact ctctgcaaa ttoagcttcc ttcccagggt ctctgtgcac tctgtcttgg 180  
 atgctctggg gagctcatgg gtggaggagt ctccaccaga gggaggctca ggggactggt 240  
 ttggccaggg atgaatattt gagggataaa aattgtgtaa gagccaaaga attggtagta 300  
 gggggagaac agagaggagc tgggctatgg gaaatgattt gaataatgga gctgggaata 360  
 ttgctggata tctgtgacta aaaaagggtc ttttaagaacc tacttcttaa tctcttcccc 420  
 aatccaaaacc atagctgtct gtccagtgct ctcttctcgc ctccagctct gccccaggct 480  
 cctcctagac tctgtcctg ggttagggca ggggaggagg gatagcaggg ttgggggaga 540  
 ggtgaggag agtgtagcat gtggggagag gaccagacct c 581

<210> 134  
 <211> 4797  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 135, 501, 4421, 4467, 4468, 4698  
 <223> n = A,T,C or G

<400> 134  
 cctgggacca aagtgtgtgc cagagctgag ggtcctggag ccacatgaga aggtctctcc 60  
 ctgtgtacct gtgcagcaca ggttaggggtg agtccactca gctgtctagg agaggaccac 120  
 ggagcagcag agacncgcca agcctttact cataccatat ttgatcctt ttccaggcaaa 180  
 ttgtggctac taatttggcc cctgaagatc aagatggctc tggggatgac tctgacaact 240  
 tctccggtct aggtgcagggt gaggttgtoa tggggggccc ccccccacaa gacggcaaca 300  
 ggtcatgcct gggggcagtg gtccaggcagt ctctgtgtt tactgagcat gtaactgagt 360  
 caccctgcct gccctgtctc caccagctgc gctccaaagg gcaatgtgca ggagaggaat 420  
 ggggtctgta gctgctgtta aggagagctc atgcttggag gtgaggtgaa ggtgtgagc 480  
 tccagaaggc cccaggggcg nctgctgcac gcaggctcat attcactagg aatagcttta 540  
 ctactaaga aacctctgga accoccttca gaaggttatt tgactctga gctctattt 600  
 tctcatctgc aaaaatggaa taataccttg acctgataag cttgtggagc tgaagagac 660  
 cacagagcca gctggggtgt agctcttcca tccaagctcc ctctcctact tccccttcc 720

tgtggggact	gggggagaga	agtccctgag	ctggagggtg	tcaggggaagc	ttcacagagg	780
aggtgtgctct	tgagtggaac	tcaggaagag	gggtgagaga	gctaaggaag	gagctctgagg	840
tcactccctgg	ggaagtgaac	tagcggaggc	ctgagagctg	caaggttagga	tactctgtgt	900
tggaaagtgtc	tggtgttgga	agtgggggcc	tttttttcag	ggaggggtgga	gccagagaag	960
tggtgtgcctc	gggataagta	ggataaacac	agtagttatg	cccctaaggg	atgcccaacc	1020
caccctctgtg	gtcacagaaa	agctttccca	ggtggcctag	gcacctgtct	cgtggctcca	1080
gagacagggt	gcacctgaca	cacacaatgg	aaggacagct	ctccttgtcc	attttccaaa	1140
gagcttagcc	tcagctgcct	tgtcagagta	ctagcctccc	tcatagcctc	agcttggcca	1200
gccacaggtgc	tctggagcct	ccccgacccc	acccaacaca	ctctgcttct	ggtcctcccc	1260
accccccaacc	tccccaacac	actctgcttc	tggtctcgca	ggtgctttgc	aagatatcac	1320
cttgtcaacag	cagacccccct	coacttggaa	ggacacgcag	ctctgcagcg	ctattccacc	1380
gtctccagaa	cccaccggcc	tggaggctac	agctgcctcc	acctccaccc	tgcggctggg	1440
agagggggccc	aaggaggagg	aggctgtagt	octggcagaa	gtggagcctg	gcctcacccg	1500
ccgggccaag	gaggccaccc	cccgaccocg	ggagaccaca	cagctcccga	ccactcatca	1560
ggcctcaacg	accacggccc	ccacggccca	ggagcccgcc	acctcccacc	cccacaggga	1620
catgcagcct	ggccaccatg	agacctcaac	ccctgcaggc	ccagcccaag	ctgacactca	1680
cactccccac	acagaggatg	gaggtccttc	tgcacccag	agggctgctg	aggatggagc	1740
ctccagctcg	ctccagcag	cagagggtcc	tggggagcag	gtgagtggcc	cttgcattcc	1800
tgggggaatt	gagtggggtg	gtctaatagc	ctggcacttg	gcaggcccta	cacctgtgcc	1860
ctgcgcgcag	tctgtattct	caccaggaag	acagggcaca	ggggcccgct	ttccctaccc	1920
ccaggggcctc	gcagagcagg	acagactaac	tatgagatca	gagcagaagc	acccttaaag	1980
atccacccaag	agaggggtccc	caaaactcaca	atccaaactt	gcagccctcg	tcgaagagtg	2040
aacgtttatc	cagtcatctt	atttatagct	tcgtgtgatt	acgctttacac	taaatagctc	2100
gctattcata	caaaatgtgt	gctttgtatc	actttttgtg	atatccatgc	catggtccag	2160
ccagggtccg	gagttgatgt	ggcaagaagg	octggctttc	ggccctgtcg	cgatccgtgt	2220
tgggtgcat	ctgagtggt	ggtggcaagg	atcaggggag	caggagctgc	ttctgggtct	2280
gtagtggagg	tggttgctgc	tgctggcggt	gacctggcca	acccaactcg	ccctggcctc	2340
ccacacggagc	tccacttttg	aaacctcggg	gggaataacg	gctgtagtgc	cgctggagcc	2400
tgaccgcggg	aaccagtcct	cagtggaatc	gggggccacg	ggggcctcac	agggcctcct	2460
ggacagagaa	gaggtctcgg	gaggtgagtt	ttctttcagg	ggggtagttt	gggtggaatt	2520
gctgctgtgg	ggtcagggtg	gggctgacca	cagccaaggc	cactgctttc	gaggggctgc	2580
ccagcagagc	caaggagccg	ctgagctgag	ctggcccgct	ctacctgcc	tgagggtcat	2640
tgcgggaggg	ctcgtggggc	tcatctttgc	tgtgtgctcg	gtgggtttca	tgctgtaccg	2700
catgaagaag	aaggacgaag	gcagctactc	cttggaggag	ccgaacaacg	ccaacggcgg	2760
ggcctaccag	aagcccccca	aacaggagga	attctatgcc	tgacgcggga	gccatgcgcc	2820
ccctcccgcc	tgccaactcac	tagggcccca	cttgctctct	ccctgaagaa	gcgcagcccc	2880
tggcctcccc	tgccaccagg	ccacctcccc	agcattccag	ccctcttggt	cgctcctgcc	2940
ccaggagtgc	tgggtgtgct	gggagctcca	ctctgcttct	ctgactctct	octggagact	3000
tagggocacca	gggggtttct	gcataggagc	tttccaccac	agccagcacc	tggtcatcgca	3060
coactctgac	tgggtttctc	caaaactgaac	cagcctctcc	ccaggtccag	ctctggagggt	3120
gaggggatgc	cgactgcttt	ggacctaaat	ggcctcatgt	ggctggaaga	tctctgggggt	3180
ggggcttggg	ctgtcacacac	ctgtagcact	tactggttag	accaagctac	ttgggggggt	3240
ggccgctggg	tggcaggagg	caggagctac	tttgtttcgt	ggggaggtct	aatcatagata	3300
tcagacttgt	tttgacatgt	tttctctag	ttctttgttc	atagccctgt	agacctgtgt	3360
acttctgagg	taagttaagt	aagttgattc	ggtatcccc	catcttggct	ccctaatacta	3420
tgttgaggag	acagcatcag	ggttaagaag	actttttttt	ttttttttaa	actagagagaa	3480
coaaactctg	aagccaaatg	ttaggcttag	tttgtgtgtt	gtctctttag	tttgtcgtct	3540
atgtgtgcaa	cagggtatgg	actatctgtc	tgggtggccc	gttctggttg	ttctgttgga	3600
ggctggccag	tccaggctgc	cgtggggccg	ccgctctttt	caagcagctg	tgccctgtgt	3660
catgcgctca	gggccaatgc	gaggcctggg	ccgctgccac	gttgagaga	ccogtgtgag	3720
aagtgaatgc	tgggactcag	ccttcagaca	gagaggactg	tagggagggg	ggcagggggc	3780
tggagctctc	octgcaggct	caoccccgct	ctcgtgtggc	gcgctctcca	ggggctgctt	3840
ccctcctggaa	attgacaggag	ggtgtcttgg	gcagagctgg	ctctgagcgc	ctccatccaa	3900
ggccaggttc	tccgttagct	octgtggccc	ccacctgggc	ccctgggctg	atacaggaat	3960
attttccaaa	gagtgatagc	cttttgcatt	tggcaaaact	ctacttaaac	caatgggttt	4020
ttccctgtac	agtagatttt	ccaaatgtta	taaaacttaa	ataagagact	ctgtgtgaat	4080
ccactgcctt	cgcttcttgc	ctctgtgtgc	tgtgtgacgt	gaccggactt	ttctgcacac	4140
accaacatgt	tgggaaactt	ggctcgaaac	tctgtgcctt	cgtctttccc	atggggaggg	4200

attctcggttc	cagggtccct	ctgtgtattt	gcttttttgt	tttgctgtaa	attctcctgg	4260
aggtcggtag	gttcagccaa	ggttttataa	ggctgatgtc	aattttctgtg	ttgcacaagt	4320
ccaagcccat	cttctaagt	gcaaagggaag	gtggatggcc	ccagcacagc	ttgacctgag	4380
gctgtgtca	cagcggagg	gtggagccga	ggcctacccc	ncagacacct	tggaatcct	4440
cctccacc	ggctgcagag	gccaganncc	agccacgggt	cctgcaactta	cttgcttatt	4500
tgacaacgtt	tcagcgactc	cgttggccac	tcagagagtg	ggccagtctg	tggaacagag	4560
atgcaccacc	aagccaaagg	aacctgtgtc	cgttatctga	tactgcgact	ttctgcctgg	4620
agtgatgcac	tgacatagc	tcgggggttg	ggaaagggtg	cggctgacca	tgtctgctgc	4680
ctggtcctgt	ggacggtnc	caagccagag	gtgggttcac	ttgtgtaacg	acaataaacg	4740
gtactgtgca	tttcgggcaa	cggctgctgt	ggtgggtgtg	gagtcctctc	ttggcct	4797

&lt;210&gt; 135

&lt;211&gt; 2856

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 135

tagtcgggg	tccccgagtg	agcacgccag	ggagcaggag	accaaaccac	gggggtcgga	60
gtcagagtcg	cagtgaggagt	ccccggaccg	gagcacagag	ctgagcggga	gagcgccgt	120
cgccagcccg	tcgccaccog	cgttaaccggc	gcagccagag	ccaccagcgc	agcgctgcca	180
tgaggccca	cagcaagaag	ctgacgggtc	gcctcatgct	ggctgtggga	ggagcagtg	240
ttggtccct	gcagtttggc	tacaacaact	gagtcataca	tgccccccag	aaggtgatcg	300
aggagttcta	caaccagaca	tggttccacc	gactatggga	gagcatcctg	cccaccacgc	360
tcaccacgct	ctggctccctc	tcagtgagca	ctttttctgt	tgggggcatg	attgtgctct	420
tctctgtggg	ccttttcggt	aaccgctttg	gcggcgagaa	ttcaatgctg	atgatgaacc	480
tgctggcctt	cgtgtccggc	gtgctcatgg	gcttctcgaa	actgggcaag	tcctttgaga	540
tgctgatcct	gggcccgttc	atcatogggt	tgtaactggg	cctgaccaca	ggcttcgtgc	600
ccatgtatgt	gggtgaagtg	tcaccacacg	cctttcgtgg	ggccctgggc	accctgcaac	660
agctggggcat	cgtcgtcggc	atctctcatg	ccacgggtgt	cggcctggag	tcocatcatg	720
gcacaagga	cctgtggccc	ctgctgctga	gcatactctc	catccggccc	ctgtgctcag	780
gcactgctgt	gccctctctg	cccagagatc	ccgccttctc	gctcatcaac	cgcaacagcg	840
agaaccgggc	caagagtggt	ctaaagaagc	tgccggggac	agctgacgtg	accatcagcc	900
tgacggagat	gaagggaagag	agtcggcaga	tgatcgggga	gaagaagggtc	accatcggtg	960
agctgttccg	ctcccccgcc	taccgcacgc	ccatcctcat	cgtgtgtgtg	ctgcagctgt	1020
cccagcagct	gtctggcact	aacgctgtct	tctattactc	cacgagcatc	ttcgagaagg	1080
cgggggtgca	gcagcctgtg	tatgccacca	ttggctccgg	tatgttcaac	acggccttca	1140
ctgtcgtgtc	gctgttttgt	gtggagcgag	caggccggcg	gcacctgcac	ctcatagggc	1200
tcgctggcat	ggcgggttgt	gccatactca	tgaccatcgc	gctagcactg	ctggagcagc	1260
taccctggat	gtcctatctg	agcatcgtgg	ccatctttgg	ctttgtggcc	ttctttgaag	1320
tggtctcgtg	cccctaccga	tggttcatcg	tggtcgaaat	cttcagccag	gttcacagtc	1380
cagctgccat	tgccgttgca	ggcttctcca	actggacctc	aaatttcaat	gtggggcatg	1440
gcttccagat	tgctggagcaa	ctgtgtggtc	ctcacgtctt	catcatcttc	actgtgctcc	1500
tggttctgtt	cttcatcttc	acctacttca	aagttcctga	gactaaaggc	cggacactcg	1560
atgagatcgc	ttccggctgc	cggcaggggg	gagccagcca	aagtgataag	acaccggagg	1620
agcgttcca	tcacctgggg	gctgatcccc	aagtgtagtt	cggccagagat	caccagcccg	1680
gcctgtctcc	agcagcccta	aggatctctc	aggagcaacg	gcagctggct	gagacttcca	1740
aacctgacac	atgtcagcgc	agccgggctc	ggggctcctt	cttcagccga	ccaatgatgt	1800
ccagaagaat	attcaggact	taacggctcc	aggattttta	caaaagcaag	actgtgtgct	1860
aaactctatto	agacaagcaa	caggttttat	aattttttta	ttactgattt	tggtattttt	1920
atatcagcct	gagtctcctg	tgcccacatc	ccaggcttca	ccctgaatgg	ttccatgctc	1980
gagggtggag	actaaagcct	gtcgagacac	ttgctctctt	caccacgcta	atctgtgaag	2040
gtgcacctat	gtctcaagga	cacactaatc	gaactatgaa	ctacaaagct	ttctatccag	2100
gaggtggcta	tgggccaccg	ttctctgtgc	ctggatctcc	ccactctagc	gggtcaggtc	2160
ctactaggat	tgccctcttc	catctcttcc	taccacaaca	ctcaaatftaa	ttttttotta	2220
cctgagacca	gttgggagca	ctggagtgca	gggaggagag	gggaaggggc	agctcgggct	2280
gccgggtctc	agtctccttt	gcactgaggg	ccacactatt	accatgagaa	gaggggcctg	2340
gggagccctg	aaactcactg	ctcaagaaga	catggagact	cctgcctcgt	tggttataga	2400
tgcaagatat	ttatatatat	ttttggttgt	caatatataa	tacagacact	aagttatagt	2460

```

atatctggac aagccaactt gtaaatacac cacctcactc ctgttaactta cctaaacaga 2520
tataaatggc tgggttttag aaacatgggt ttgaaatgct tgtggattga gggtaggagg 2580
tttggatggg agtgagacag aagtaagtgg ggttgcaacc actgcaacgg cttagacttc 2640
gactcaggat ccagtcctt acacgtaccc ctcatcagtg tctctttgct caaaaatctg 2700
tttgatccct gttaccaga gaatatatac attctttatc ttgacattca aggcatttct 2760
atcacatatt tgatagtgg tgttcaaaaa aacactagtt ttgtgcacg cgtgatgctc 2820
aggttgtaaa tcgcattatt ttgaatgtga agggaa 2856

```

```

<210> 136
<211> 356
<212> DNA
<213> Homo sapiens

```

```

<400> 136
ggtggagcca aatgaagaaa atgaagatga aagagacaga cacctcagtt tttctggatc 60
aggcattgat gatgatgaag attttatctc cagcacatt tcaaccacac cccgggcttt 120
tgaccacaca aaacagaacc aggactggac tcagtggaa ccaagccatt caaatccgga 180
agtgtcactt cagacaacca caaggatgac tgatgtagac agaaatggca ccaactgctta 240
tgaaggaaac tggaaaccag aagcacaccc tccctcatt caccatgagc atcatgagga 300
agaagagacc ccacattcta caagcacaat ccaggcaact cctagtagta caacgg 356

```

```

<210> 137
<211> 356
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 254, 264, 279, 281, 290, 328, 342
<223> n = A,T,C or G

```

```

<400> 137
gcagggtggag aagacathtt attgttctg ggtctctgagg aggccattg gtggggctg 60
gtcactggct gcccccggaa caggcgctg ctccatggct ctgcttggg tagtctggtg 120
ctatgtctcc cagcaaggac agaaactcag aaaaatcaat cttcttatcc tcattcttgt 180
cctttttctc aaagacatcg gcgaggtaat ttgtgccctt tttaacctcg cccgcgacca 240
cgctaaggcc aaanttcag acanayggcc gggcgggtnc nataggggan cccaacttgg 300
ggaccacaac tctggcgcgg aaacacangg gcataagett gnttctctg gggaaa 356

```

```

<210> 138
<211> 353
<212> DNA
<213> Homo sapiens

```

```

<400> 138
aggtccagtc ctccaattgg cctgatgaga gtgggggagt gcaagggagc tttctctgc 60
aatagacact tagattttctc tcttgtggga agaaaccacc tgtccatcca ctgactcttc 120
tacattgatg tggaaattgc tgctgtacc accacctcct gaagaggctt cctgatgccc 180
aatgccagcc atcttggcat cctggccctc gagcaggctg cggtaagtga cgatctctg 240
ctccagcggt gtctttatgt caagcagcat cttgtactcc tggttctgag cctccatctc 300
gcacgcggagc tcactcagac ctgscggsg mssmcgctam gccgaattcc agc 353

```

```

<210> 139
<211> 371
<212> DNA
<213> Homo sapiens

```

```

<400> 139

```

```

agcgtgtgctg cggcccgaggt ccatccgaag caagattgca gatggcagtg tgaagagaga 60
agacatattc tacacttcaa agcttttggt caattcccat cgaccagagt tgggtccgacc 120
agccttggaag aggtcactga aaaatcttca attggattat gttgacctct accttattca 180
ttttccagtg tctgtaaaag caggtgagga agtgatccca aaagatgaaa atggaaaaat 240
actatttgac acagtggtatc tctgtgccac gtgggaggcc gtggagaagt gtaaaagtgc 300
aggattggac ctgcccgggc ggccgctcga aagccgaatt ccagcacact ggcggccgtt 360
actagtggat c 371

```

```

<210> 140
<211> 370
<212> DNA
<213> Homo sapiens

```

```

<400> 140
tagcgtggtc gggcccgaggt tccatctccc tttgggaact agggggctgc tgggtgggaaa 60
tgggagccag ggcagatggt gcatttccttt gtgtccctgt aaatgtggga ctacaagaag 120
aggagctgcc tgagtgtgtac tttctcttcc tggtaatctc ctggcccgag ctcactggcag 180
aatagaggtta tttttaggct atttttgtaa tatggcttct tatggccttc cctgtgtagc 240
tgaattccca agccctgcat tgtacagccc ccactcccc tcaccaccta ataaaggaaat 300
agttaacact caaaaaaaaa aaaaacacct cccggggcggc gcgtcgaaag ccgaattcca 360
gcacactggc 370

```

```

<210> 141
<211> 371
<212> DNA
<213> Homo sapiens

```

```

<400> 141
tagcgtggtc gggcccgaggt tccctctgtgc tgcctgtcac agcccgatgg taccagcgca 60
gggtgtaggc agtgccaggag cctccatcca gtggcaggga acaggggtca tcaactatccc 120
aaggagcttc aggttctctg taactctcca cagaatactc ggagtattca gagtactcat 180
catctcagg gggtaaccgc tcttctctct ctgcatgaga gacggcgagc acaggcacag 240
catggagctg ggagccggca gtgtctcgag cataactagg gagggtgct gatccagatg 300
cgatgaactg gccctggcag gcacagtgtc gactcatctc ttggcgacct gccccggcgg 360
ccgctcgaa g 371

```

```

<210> 142
<211> 343
<212> DNA
<213> Homo sapiens

```

```

<400> 142
gcgttttgag gccaatggtg taaaaggaaa tatcttcaca taaaaactag atggaagcat 60
tgtcagaaac ctcttttgta tgttttgctt caactcacag agttgaacat tctttttcat 120
agagcagttt tgaacaactc ttttgtagaa ttgccaagcg gatgattgga tcgctatgag 180
gtcttcattg gaaacgggat acctttacat aaaaactaga cagttagcatt ctcagaaatt 240
tctttgggat gtgggcatte aaccacaga ggagaacttc atttgataga gcagttttga 300
aacacccttt ttgtagaatc tacaggttga cattttagtg gct 343

```

```

<210> 143
<211> 354
<212> DNA
<213> Homo sapiens

```

```

<400> 143
aggctctgat gcagaaaaac tcagactgtc tgcaacttta cagatgggtgc attggttcag 60
catcaggagt gggatgggaa ggaaagcaca ataacaagaa aattgaaaga tgggaaatta 120
gtgggtggagt gtgtcatgaa caatgtcacc tgtactcgga tctatgaaaa agtagaataa 180

```

aaattccatc	atcacttttg	acaggagtta	attaagagaa	tgaccaagct	cagttcaatg	240
agc aaaatctc	catactgttt	ctttcttttt	tttttcatta	ctgtgttcaa	ttatctttat	300
cataaacatt	ttacatgcag	ctatttcaaa	gtgtgttgga	ttaatttagga	tcat	354

&lt;210&gt; 144

&lt;211&gt; 353

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 144

ggtcaaggac	ctgggggacc	cccaggtcca	gcagccacat	gattctgcag	cagacaggga	60
cctagagcac	atctggatct	cagcccccac	cctggcaacc	tgccctgccta	gagaactccc	120
aagatgacag	actaagttagg	attctgccat	ttagaataat	tctggatccc	tgggcggtgc	180
gttaagtgtc	ttacttttca	ttctgtctta	cgatagtctt	cagaggtggg	aacagatgaa	240
gaaaccatgc	cccagagaag	gttaagtgc	ttctctotta	tgaggccagt	gttccaaact	300
agggtttgct	gataccagac	ctgtggcccc	acctcccatg	cagggtctctg	tgg	353

&lt;210&gt; 145

&lt;211&gt; 371

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 145

caggctctgtc	ataaactggt	ctggagtctc	tgacgactcc	ttgttoacca	aatgcaccat	60
ttctctgagac	ttgctggcct	ctccgttgag	tccaacttggc	ttctgtctct	ccacagctcc	120
attgccactg	ttgatcacta	gctttttctt	ctgcccacac	cttcttcgac	tgtttgactgc	180
aatgcaaaact	gcaagaatca	aagccaaggc	caagagggat	gccaagatga	tcagccattc	240
tggaatttgg	gggtgtcotta	taggaccaga	ggttgtgttt	gtcccaactt	cttgactccc	300
atgtgagacc	tgcggcgcga	ccacgcctaag	ccgaattcca	gcacactggc	ggcccgctac	360
tagtggatcc	g					371

&lt;210&gt; 146

&lt;211&gt; 355

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 146

ggtctctccgt	cctcttccca	gaggtgtcgg	ggcttggccc	cagctctccat	cttctgtctct	60
caggatggcg	agtgcagcgg	gctccaaggc	tgaattcatt	gtcggaggga	aatataaaact	120
ggtagcggaag	atcggtctcg	gctccttcgg	ggacacttat	ttggcgatca	acatcaccaa	180
cgggcgaggaa	gtggcagtg	agctagaatc	tcagaaggcc	aggcatcccc	agttgtgtga	240
cgagagcaag	ctctataaga	ttcttcaagg	tggggttggc	atccccca	tacgggtgga	300
tggtcaggaa	aaagactaca	atgtactagt	catggatctt	ctgggacctc	gcctc	355

&lt;210&gt; 147

&lt;211&gt; 355

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 147

ggtctgtttac	aaaatgaaga	cagacaacac	aacatttact	ctgtggagat	atcctactca	60
tactatgcac	gtgctgtgat	tttgaacata	actcgtccca	aaaacttgc	acgatcatcc	120
tgactttttta	ggttgctcta	tcctcaatc	ttgcaactca	ctgttacttc	tttcccagtg	180
ttgttaggag	caaagctgac	ctgaacagca	accaatggct	gtagataccc	aacatgcagt	240
tttttcccat	aatatgggaa	atatttttaag	tctatcttc	cattatgagg	ataaactgct	300
acatttggta	tatcttcatt	ctttgaaaca	caatctatcc	ttggcaactcc	ttcag	355

&lt;210&gt; 148

<211> 369  
 <212> DNA  
 <213> Homo sapiens

<400> 148  
 aggtctctct cccctctctc ctctcctgcc agccaagtga agacatgctt aettcccttt 60  
 caccttctct catgatgtgg gaagagtgtc gcaaccacgc cctagccaac accgatgag 120  
 agggagtggt cgagggtctt ctgagaaggt ttctctcaca cttagaaga agcgcttaag 180  
 atgtggcagc cctctctctt caagtggctc ttgtcctgtt gccctgggag ttctcaaat 240  
 gctgcagcag cctccatcca gcoctgaggat gacatcaata cacagaggaa gaagagtcat 300  
 gaaaagatga gagaagttac agactctcct gggcgacccc gagagcttac cattctctcag 360  
 actttctca 369

<210> 149  
 <211> 620  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 169, 171, 222, 472, 528, 559, 599  
 <223> n = A,T,C or G

<400> 149  
 actagtcaaa aatgctaaaa taatttggga gaaaaatatt tttaagtagt gttatatgtt 60  
 catgtttatc ttttattatg ttttgtgaag ttgtgtcttt tcaactaatta cctatactat 120  
 gccaatattt cttatatctt atccataaca ttatactact atttgaana naatatgac 180  
 gtgaaactta acactttata aggtaaaaat gaggtttcca anatttaata atctgatcaa 240  
 gttcttggta ttccaataa gaatggactt ggtctgttaa gggctaagga gaagaggaa 300  
 ataagggttaa aagtgtgtta tgaccaaaca ttctaaaaga aatgcaaaaa aaaagtttat 360  
 ttccaagcct tgaactatt taaggaaaagc aaaatcattt cctaaatgca tatcatttgt 420  
 gagaatttct cattaatact ctgaatcatt catttcaata aggcctatgt tnaactccgat 480  
 atgtctctaa gaaagtacta ttctatgtgc caaacctggt tgccatantt gggtaaggc 540  
 ttctccttaa gtgtgaaant atttaaaatg aaattttcct ctttttaaaa attctttana 600  
 aggggttaag gtgttgggga 620

<210> 150  
 <211> 371  
 <212> DNA  
 <213> Homo sapiens

<400> 150  
 ggtccgatca aaacctgcta cctccccaag actttactag tgccgataaa ctttctcaaa 60  
 gagcaaccgc tatcaacttc ctgtttataa aacctctaac catctctttg ttctttgaa 120  
 atgctgaaaa ccacctggct tgcatgtatg ccgaatttg yaattctttt ctctcaaatg 180  
 aaaaatttaa tttagggatt catttctata ttttcacata tgtagtatta ttatttctct 240  
 atatgtgtta ggtgaaattt atggtatttg agtggtcgaag aaaatatatt tttaaagctt 300  
 tcatttttcc ccagtggaat gatttagaat tttttatgta aatatacaga atgttttttc 360  
 ttacttttat a 371

<210> 151  
 <211> 4655  
 <212> DNA  
 <213> Homo sapiens

<400> 151  
 gggacttgag ttctgttata ttcttaagta gattcatatt gtaagggtct cggggtgggg 60  
 ggggtggcaa aatctggag ccagaagaaa ggacagcagc attgatcaat cttacagcta 120



acatgttgta	cctggaaaac	aatgcccgaga	ctcaatttag	tgagccacag	tacacgaacc	180
tggggctcct	gaacagcatg	gaaccagcaga	tccagaacgg	ctcctcgtcc	accagctccct	240
ataacacaga	ccacgcgcag	aacagcgcga	cggcgccctc	gcocctacgca	cagcccgagct	300
ccaccttcga	tgtctctctc	ccatcaccgc	ccatccctcc	caacaccgac	taccagggcc	360
cgccagcttt	cgacgtgtcc	ttccagcagt	cgagcacccg	caagtccggc	acctggagct	420
attccactga	actgaagaaa	ctctactgcc	aaattgcaaa	gacatgcccc	atccagatca	480
aggtgatgac	ccacacctct	caggggagctg	ttactccggc	catgcctgtc	ctcaaaaaag	540
ctgagcagct	cacggaggtg	gtgaagcggg	gccccaaoca	tgagctgagc	tgtaatttca	600
acogagggaca	gattggccct	ycatgtcatt	tgattcagat	agaggggaac	agccatgcc	660
agtatgtaga	agatcccatc	acaggaagac	agagtgtgct	ggtaacctat	gagccaaccc	720
aggttggcac	tgaattcacg	acagtcttgt	acaatttcac	gtgtaacacg	agttgtgttg	780
gagggatgaa	ccgcgctcca	attttaatca	ttgttactct	ggaaccacga	gatggggcaag	840
tcctgggccc	acgctgcttt	gaggcccgga	ctctgtgctg	cccgaggaaga	gacaggaagg	900
cggtatgaaga	tagcatcaga	aagcagcaag	tttcggacag	tacaaagaac	ggtgatggta	960
cgaaagccccc	gtttcgtcag	aacacacatg	gtatccagat	gacatcccat	aagaaacgaa	1020
gatcccgaca	tgttggaact	gtatactttac	cagtggaggg	catgtgagact	tatgaattgc	1080
tggtgaagat	caaaagctcc	ctggaaactca	tgacgtacct	tcttcagcac	acaattgaaa	1140
cgatggcgac	acagcaaacg	cagcagcacc	agcaattact	tcagaaaaacg	acctcaatca	1200
agtcctccatc	ttcatatggt	aacagctccc	cacctctgaa	caaatgaac	agcatgaaca	1260
agctgccttc	tgtgagccag	cttatcaacc	ctcagcagcg	caacgccctc	actcctacaa	1320
ccattctctga	tggatgggga	gccaacatcc	ccatgatggg	caaccacatg	ccaatggctg	1380
gagacgatgaa	tggactcagc	cccaccagg	cactccctcc	cccaactctc	atggcatcca	1440
ctccccaact	cacaccocca	cctccgttat	ccacagattg	cagcattgtc	agttctctag	1500
cgaggttggg	ctgtttcatca	tgtctggact	atttcacgac	ccaggggctg	accacacatc	1560
atcagattga	cgattactcc	atggatgac	tggcaagtct	ggaataccct	gagcaatttc	1620
gacatcgcat	ctggaagggc	atcctggacc	accggcagct	ccacgaattc	tctctccctt	1680
ctcatctcct	gcggaaccca	agcagtgctc	ctacagtacg	tgtgggctcc	agtgagagcc	1740
ggggtgagcg	tgttattgat	gctgtgcgat	tcacctccg	ccagaccatc	tcttccacc	1800
cccgagatga	gtggaatgac	ttcaactttg	acatggatgc	tcggccgcaat	aagcaaacgc	1860
gcatcaaga	ggagggggag	tgagcctcac	catgtgagct	cttctctacc	ctctcctaac	1920
tgccagcccc	ctaaaagcac	tcctgcttaa	tcttcaaaag	cttctcccta	gctcctcccc	1980
ttctctctgt	ctgatttctt	aggggaagga	gaagtaagag	gcttactctt	taccctaaac	2040
atctgacctg	gcactctaatt	ctgattctgg	ctttaagctc	tcaaaactat	agcttgcaag	2100
actgtagctt	gcatggtcta	ggtagaagtg	agcaaaaaag	agttgggtgt	ctccttaagc	2160
tgcagagatt	tctcattgac	ttttataaag	catggtcac	cttatagctt	aagactatat	2220
atataaatgt	ataaatatac	agtatagatt	tttgggtggg	gggcattgag	tattgtttaa	2280
aatgttaatt	aaatgaaaga	aaattgagtt	gcacttattg	accatttttt	aatttactgt	2340
ttttggatgg	cttgtctata	ctcctccctc	taaggggtat	catgtatggt	gatagggtac	2400
tagagcttaa	tgtcacatgt	gagtgacgat	gagtgacaga	ttctttcagt	cttttggatt	2460
ctaaatcatg	gccacatcaa	acotttgagt	agatccattt	ccattgtcat	ttatgtgcat	2520
aagactgtag	atatgtatcc	ttttctcagt	gttggtatat	tttatattac	tgacatttct	2580
ctcagtgtatg	atggttccag	ttggggtaga	ttaatccagt	tataagaaga	agttctatgc	2640
caaacgctct	ctttagtttt	tggttgggaa	tgaggaaaat	tcttaaaaag	cccatagcag	2700
ccagttcaaa	aacaccgcag	gtcatgtatt	tgagcatcat	agtaaccocc	ttaaaatttaa	2760
tccagataac	ttatatctac	aatattgcat	ggagaggaat	ttgtgcgcat	tacagaggta	2820
ttaaaaactaa	atttccatca	tagattgact	aaactaaata	caactttgct	actgttgtaga	2880
gtaattctgat	tgtattgtat	gggatgaatt	ccatctatct	ctgtataaca	aatgtgcatc	2940
actgtctatt	aatattccag	gtaaatagga	atcattcaga	aatgtttgag	ctgtactaaa	3000
cagtaagata	tctcaatgaa	ccataaatcc	aaactttgtaa	aaactttggt	aagcatagat	3060
aatattgttt	ggtaaatggt	tcttttgttt	ggtaaatggt	tcytttaaag	acctcctcat	3120
ttctataaaac	ctcgtcatgta	gaggcttgtt	tacotttctc	tctctaaagt	ttacaagtgt	3180
agtggtgtatt	tgaaaaatatt	aaaattatga	tactgtgttt	ccctgaggtt	aatatgcatc	3240
actgtatctt	ttttcttttt	aaccggtaag	agtttcagtt	tgttggaaga	taactctgag	3300
aaccctgttt	cccgctccat	tcctctagg	actaccocata	gacatgagga	gtccccacag	3360
agcaagagat	aagtctttca	tggctgctgt	tgcttaaac	acttaaacga	agagttccct	3420
tgaaaacttt	ggaaaaatcg	ttaatgacaa	tattccagat	ctttcagaaa	tataaacatc	3480
ttttttgcgt	gcattgcaaat	gagctctgaa	tattcccat	gcattctggt	acgggctgt	3540
cattgacat	aagcttccat	tttaatttta	aagtgcaaaa	gggcagcgt	ggctctaaaa	3600

```

ggtaagtgtg ggattgcctc tgaaaagtgt gtatatattt tgtgtgaaat tgcatacttt 3660
gtatttttgt tatttttttt ttctttcttg gatagtggga ttccocagaac cacacttgaa 3720
accctttttt atcgtttttt tattttcatg aaaataccat tttagtaagaa taccacatca 3780
aataagaaat aatgctacaa ttttaagagg ggagggaagg gaaagttttt ttttttatta 3840
tttttttaaa atttttgatg ttaaagagaa tgagtccttg atttcaaagt tttgttgtag 3900
ttaaatggta ataagcactg taaacttctg caacaagcat gcagctttgc aaaccattta 3960
aggggaagaa tgaagactgt tccttggtcc tagtaagaag acaaactgct tcccttactt 4020
tgcgtagggt ttgaataaac ctaggacttc cgagctatgt cagtactatt caggtaaacac 4080
tagggccttg gaaatccctg tactgtgtct catggatttg gcactagcca aagcgaggca 4140
ccocttactg gottacctcc tcatggcagc ctactctcct tgagtgtatg agtagccagg 4200
gtaaggggta aaaggatagt aagcatagaa accactagaa agtggggcta atggagtctt 4260
tgtggccctca gctcaatgca gttagctgaa gaattgaaaa gtttttggtt ggagacgttt 4320
ataaacagaa atggaagaca gagttttcat taaatccttt tacctttttt ttttcttggt 4380
aatccccctaa aataacagta tgtgggatat tgaatgttaa aggggatattt ttttctatta 4440
tttttataat tgtacaaaat taagcaaatg ttaaaagtgt tatatgcttt attaatgttt 4500
tcaaaaggtta ttatacatgt gatacatttt ttaagcttca gttgcttgct ttctgggtac 4560
ttctgttatg ggcttttggg gagccagaag ccaatctaca atctcttttt gtttgcagg 4620
acatgcaata aaatttaaaa aataaataaa aacta 4655

```

&lt;210&gt; 152

&lt;211&gt; 586

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 152

```

Met Leu Tyr Leu Glu Asn Asn Ala Gln Thr Gln Phe Ser Glu Pro Gln
1 5 10 15
Tyr Thr Asn Leu Gly Leu Leu Asn Ser Met Asp Gln Gln Ile Gln Asn
20 25 30
Gly Ser Ser Ser Thr Ser Pro Tyr Asn Thr Asp His Ala Gln Asn Ser
35 40 45
Val Thr Ala Pro Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Ala
50 55 60
Leu Ser Pro Ser Pro Ala Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro
65 70 75 80
His Ser Phe Asp Val Ser Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala
85 90 95
Thr Trp Thr Tyr Ser Thr Glu Leu Lys Lys Leu Tyr Cys Gln Ile Ala
100 105 110
Lys Thr Cys Pro Ile Gln Ile Lys Val Met Thr Pro Pro Pro Gln Gly
115 120 125
Ala Val Ile Arg Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr
130 135 140
Glu Val Val Lys Arg Cys Pro Asn His Glu Leu Ser Arg Glu Phe Asn
145 150 155 160
Glu Gly Gln Ile Ala Pro Ser Ser His Leu Ile Arg Val Glu Gly Asn
165 170 175
Ser His Ala Gln Tyr Val Glu Asp Pro Ile Thr Gly Arg Gln Ser Val
180 185 190
Leu Val Pro Tyr Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Val
195 200 205
Leu Tyr Asn Phe Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg
210 215 220
Arg Pro Ile Leu Ile Ile Val Thr Leu Glu Thr Arg Asp Gly Gln Val
225 230 235 240
Leu Gly Arg Arg Cys Phe Glu Ala Arg Ile Cys Ala Cys Pro Gly Arg
245 250 255
Asp Arg Lys Ala Asp Glu Asp Ser Ile Arg Lys Gln Gln Val Ser Asp

```

Ser	Thr	Lys	Asn	Gly	Asp		Gly	Thr	Leu	Arg	Pro	Phe	Arg	Gln	Asn	Thr
His	Glu	Ile	Gln	Met	Thr		Ser	Ile	Lys	Lys	Arg	Arg	Ser	Pro	Asp	Asp
Glu	Val	Tyr	Leu	Pro			Val	Arg	Gly	Arg	Glu	Thr	Tyr	Glu	Met	Leu
Val	Lys	Ile	Lys	Glu	Ser		Leu	Glu	Leu	Met	Gln	Tyr	Leu	Leu	Gln	His
Thr	Ile	Glu	Thr	Tyr	Arg		Gln	Gln	Gln	Gln	Gln	Gln	His	Gln	His	Leu
Leu	Gln	Lys	Gln	Thr	Ser		Ile	Gln	Ser	Pro	Ser	Ser	Tyr	Gly	Asn	Ser
Ser	Pro	Pro	Leu	Asn	Lys		Met	Asn	Ser	Met	Asn	Lys	Leu	Pro	Ser	Val
Ser	Gln	Leu	Ile	Asn	Pro		Gln	Gln	Arg	Asn	Ala	Leu	Thr	Pro	Thr	Thr
Ile	Pro	Asp	Gly	Met	Gly		Ala	Asn	Ile	Pro	Met	Met	Gly	Thr	His	Met
Pro	Met	Ala	Gly	Asp	Met		Asn	Gly	Leu	Ser	Pro	Thr	Gln	Ala	Leu	Pro
Pro	Pro	Leu	Ser	Met	Pro		Ser	Thr	Ser	His	Cys	Thr	Pro	Pro	Pro	Pro
Tyr	Pro	Thr	Asp	Cys	Ser		Ile	Val	Ser	Phe	Leu	Ala	Arg	Leu	Gly	Cys
Ser	Ser	Cys	Leu	Asp	Tyr		Phe	Thr	Thr	Gln	Gly	Leu	Thr	Thr	Ile	Tyr
Gln	Ile	Glu	His	Tyr	Ser		Met	Asp	Asp	Leu	Ala	Ser	Leu	Lys	Ile	Pro
Glu	Gln	Phe	Arg	His	Ala		Ile	Trp	Lys	Gly	Ile	Leu	Asp	His	Arg	Gln
Leu	His	Glu	Phe	Ser	Ser		Pro	Ser	His	Leu	Leu	Arg	Thr	Pro	Ser	Ser
Ala	Ser	Thr	Val	Ser	Val		Gly	Ser	Ser	Glu	Thr	Arg	Gly	Glu	Arg	Val
Ile	Asp	Ala	Val	Arg	Phe		Thr	Leu	Arg	Gln	Thr	Ile	Ser	Phe	Pro	Pro
Arg	Asp	Glu	Trp	Asn	Asp		Phe	Asn	Phe	Asp	Met	Asp	Ala	Arg	Arg	Asn
Lys	Gln	Gln	Arg	Ile	Lys		Glu	Glu	Gly	Glu						

<210> 153

<211> 2007

<212> DNA

<213> Homo sapiens

<400> 153

gaattgctgc	ctgctcagg	gaaagttctg	ttactccact	gactctctct	tttctgtgta	60
caattgcgcg	caagaagaat	attacatgct	tttgagcaac	aggagctcaa	ggttggtctc	120
tggccagggc	aattttggag	agcaaaaaat	ttgcagctgag	acagctgacc	aggggtgtga	180
cttgacaaa	tgccttgctg	ctccagcgcc	ttgagctgtc	ggtgtctaaa	ggtgacctga	240
atgtataaac	atcggtggac	agtgctctaa	agtggtgtcta	ttgggacctt	tgtgtgacca	300
actctctgga	ccctctcaac	caagataaag	aagttgtctgc	ggggaaagct	gtggcagact	360
cgcgcacaag	cttggtgtct	aagacagctg	ttcacagcgt	ctgcagagac	ctcaagcgac	420
tgcagctgtg	caagctgtgc	gtccgcagct	ttgcacagca	ggcgactgtg	gagggaacta	480
ttctggtcat	tgcgtatccc	atdaccagtc	tccacatcag	ggccctagtt	gaaagtcttc	540

tgcggcgctg	gcgcccgctg	aaagcctctg	atggagatta	ctacaccttg	gctgtaccga	600
tgggagatgt	accaatggat	ggtatctctg	ttgctgatat	tggagcagcc	gtctctagca	660
tttttaattc	tccagaggaa	tttttaggca	aggccgtggg	gctcagtgca	gaagcactaa	720
caatcacgca	atatgctgat	gttttgcaca	aggctttggg	gaaagaagtc	cgagatgcaa	780
agattacccc	ggaagcttct	gagaagctgg	gattccctgc	agcaaaggaa	atagccaata	840
tgtgtcgttt	ctatgaaatg	aagccagacc	gagatgtcaa	tctcacccac	caactaaatc	900
ccaaagtcag	aagcttcagc	cagtttatct	cagagaacca	gggagccttc	aaggcagatg	960
agaaaatcag	ctgttcagat	aggcctctgc	accacacagc	ctcttctctc	ctctgatcctt	1020
ttcctcttta	cggcacacaa	ttcatgttga	cagaacatgc	tggaaatgcaa	ttgttttgcaa	1080
caccgaagga	tttctcgctg	tgcctctctc	agttaggaagc	actgcattgg	tgataggaca	1140
cggttaattgt	attcacattt	aacttgctag	ttagtgtata	gggtggtagca	actgttttgtt	1200
aaaaatgagaa	gcttcgggaa	ttggagcttc	tctctaccac	ctaattgggag	ggcagatttat	1260
actgggattt	ctcctgggtg	agtaatttca	agccctaagt	ctgaaattcc	cttagcgacg	1320
tccagttttc	tcaactgcac	tgcacaaattc	ccagtgaact	tttaagtact	tttaacttaa	1380
aaaaatgaaac	atctttgtag	agaattttct	ggggaacatg	gtgtccaagt	aacaagcaca	1440
agcattggaa	atgctaaaaat	tcagttttgc	ctcaagattg	gaagttttatt	ttctgctgaca	1500
ttcatgaagt	catctatgta	gccacacatc	aatttatcat	ctattaattc	cttgatcctt	1560
cattttacga	ttctgcaaac	ttttcttgag	caccagcacg	gggtggcatt	tgtggaacttc	1620
tcttcattcc	tatgtgtttt	cttatcaaa	tgtatccact	tgcgaagcct	cctttccagt	1680
ctgtggttgg	gttcaagtca	tgcacgggac	agggggccca	tctctctgtt	tgctcttagg	1740
caaaaatccag	gggatctgca	gtggggagga	ggggcaggaa	gctggagcga	agccctgtga	1800
agggtagggga	tgtggaaaga	caaggtgaca	gaaggaccac	ataggacott	tctatatctc	1860
tggcttagca	ttttctacac	catattgtaa	tgtctttatt	tgtgagtttt	cttcgactag	1920
gtgagtgact	aacagtcact	tttatccacg	tgcctgtgac	ataataagtg	atcaataaat	1980
gttgattgac	taaaaaaaaa	aaaaaaaa				2007

&lt;210&gt; 154

&lt;211&gt; 2148

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 154

gaattcgtcg	ctgctccagg	gaaagttctg	ttactccact	gactctctct	tttctcgata	60
acatggccag	caagaaagta	attacagtg	ttggagcaac	aggagctcaa	ggtggctctg	120
tggccagggc	aattttggag	agcaaaaaat	ttgcagtgag	agcagtgacc	agggatgtga	180
cttgaccaaa	tgcctctggg	ctccagcgcc	ttggagctga	ggtggtcaaa	ggtgacctga	240
atgataaagc	atcgtgtggc	agtgccctaa	aaggggaaagc	tggtggcaga	cttcgccaa	300
cacotgggtc	tgaagcacgt	ggtgtacagc	ggcctggaga	acgtcaagcg	actgacggat	360
ggcaagctgg	aggtgcgcga	ctttgacagc	aaggggcagg	tggagggagta	cttctggctgc	420
attggctacc	ccatgaccac	tgtccgcgtg	ctggccactc	tgaaaaaact	tctcgcggcg	480
tggcgccccc	tgaagacccc	tgtatggagat	tactacacot	tggctgtacc	gatggggagat	540
gtaccaatcg	atgggtatct	tgttgcgtgat	attggagcag	cgcctctctag	catttttaat	600
tctccagagg	aatttttagg	caaggccgtg	gggctcagtg	cagaagcact	aacaatacac	660
caatatatcg	atgttttctc	caaggctttg	gggaagaag	tccgagatgc	aaagactatc	720
tgtgctatag	atgacacagaa	aacagtgga	gaaggtttca	tggaaagcgt	gggcttgagt	780
tgtgcttgta	gggaacatga	ccatgtatag	acagaggagg	catcaagaag	gctgggctgt	840
ctaattctcg	aataaacacg	acaaacacga	gcagctacgg	gaaggaggca	aattctggct	900
tgcctcttat	ccttgattac	cccggaaagct	ttcgagaagc	tgggattccc	tgcacaaagc	960
gaaatagcca	atatgtgtcg	tttctatgaa	atgaagccag	accgagatgt	gacatctaac	1020
caaccaactaa	atcccaaatg	caaaagcttc	agccatttta	tctcagagaa	ccaggggagcc	1080
tctcagggca	tgtagaaaaat	cagctgttca	gataggcctc	tgcacccacac	agcctctttc	1140
ttctatgatc	ctttctctct	ttacggcaca	acattcagtg	tgcacagaca	gaattctggat	1200
caattgtttg	caacacccga	ggatttctgt	cggctgcctc	ttcagtagga	agcactcgat	1260
tgtgttagtg	acacggtaat	ttgattcaca	attgacttgc	tagtttagtga	taagggtgggt	1320
acaactgttt	ggtaaaatga	gaagcctcgg	aacttggagc	ttctctccta	ccaataatgg	1380
gagggcagat	tatactggga	tttctctctg	gtgagtaatt	tcaagcccta	atgctgaaat	1440
tccctcagtg	atgctcagtt	ttctcaactg	cattgcacaaa	ttccagtgta	acttttaagt	1500
acttttaact	taaaaaaaaa	aacatctttg	tagagaattt	tctgggggaa	atggtgttca	1560

```

atgaacaagc acaagcattg gaaatgctaa aattcagttt tgcctcaaga ttggaagttt 1620
atttttctgac tcattcatga agtcattctat tgagccacca ttcaattatt catctattaa 1680
ttccttgatc ctctcatttat ccattctgca aacttttctt gagcaccacag acgggtggcc 1740
atttgtggag ttctcttcat tctctatgtt tttttatca aagtgatcca ctctcgaaag 1800
gctcctttcc agtctgtggt tgggttcaag tcatgccaag gccagggggg ccactctctc 1860
gttttagctct agggcaaatc caggggatct gcagtgggga gcgggggcag gaagctggag 1920
ggaaggcctg tgaagggtag ggaatgtgaa agacaaggtg acagaaggac ccaataggac 1980
ctttctatat ctctggctta gcaattttcta catcatattg taatcgctct atttgctagt 2040
tttcttctct actgtgagtg actaacagtc atctttatcc cagtgctctg tacataataa 2100
gtgatcaata aatgttgatt gactaaatga aaaaaaaaaa aaaaaaaaaa 2148

```

&lt;210&gt; 155

&lt;211&gt; 153

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 155

```

Met Thr Ser Val Arg Val Ala Ala Tyr Phe Glu Asn Phe Leu Ala Ala
 1           5           10          15
Trp Arg Pro Val Lys Ala Ser Asp Gly Asp Tyr Tyr Thr Leu Ala Val
      20          25
Pro Met Gly Asp Val Pro Met Asp Gly Ile Ser Val Ala Asp Ile Gly
      35          40          45
Ala Ala Val Ser Ser Ile Phe Asn Ser Pro Glu Glu Phe Leu Gly Lys
      50          55          60
Ala Val Gly Leu Ser Ala Glu Ala Leu Thr Ile Gln Gln Tyr Ala Asp
      65          70          75          80
Val Leu Ser Lys Ala Leu Gly Lys Glu Val Arg Asp Ala Lys Ile Thr
      85          90          95
Pro Glu Ala Phe Glu Lys Leu Gly Phe Pro Ala Ala Lys Glu Ile Ala
      100         105         110
Asn Met Cys Arg Phe Tyr Glu Met Lys Pro Asp Arg Asp Val Asn Leu
      115         120         125
Thr His Gln Leu Asn Pro Lys Val Lys Ser Phe Ser Gln Phe Ile Ser
      130         135         140
Glu Asn Gln Gly Ala Phe Lys Gly Met
145           150

```

&lt;210&gt; 156

&lt;211&gt; 128

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 156

```

Met Thr Ser Val Arg Val Ala Ala Tyr Phe Glu Asn Phe Leu Ala Ala
 1           5           10          15
Trp Arg Pro Val Lys Ala Ser Asp Gly Asp Tyr Tyr Thr Leu Ala Val
      20          25
Pro Met Gly Asp Val Pro Met Asp Gly Ile Ser Val Ala Asp Ile Gly
      35          40          45
Ala Ala Val Ser Ser Ile Phe Asn Ser Pro Glu Glu Phe Leu Gly Lys
      50          55          60
Ala Val Gly Leu Ser Ala Glu Ala Leu Thr Ile Gln Gln Tyr Ala Asp
      65          70          75          80
Val Leu Ser Lys Ala Leu Gly Lys Glu Val Arg Asp Ala Lys Thr Ile
      85          90          95
Cys Ala Ile Asp Asp Gln Lys Thr Val Glu Glu Gly Phe Met Glu Asp

```

	100		105		110	
Val Gly Leu Ser Trp Ser Leu Arg Glu His Asp His Val Ala Gly Ala						
	115		120		125	
<210>	157					
<211>	424					
<212>	DNA					
<213>	Homo sapiens					
<220>						
<221>	misc_feature					
<222>	320, 322					
<223>	n = A,T,C or G					
<400>	157					
ctgcagccgc	ggggatccac	tagtccagtg	tggtggaatt	catttggtctt	tacaagactt	60
ggatcacatta	cagcagacat	ggaatatata	ttttaaaaaa	tttctctcca	acctcttcca	120
aattcagtc	ccactgttat	attacctctt	ccaggaaccc	tccagtgggg	aaggctgcga	180
tattagattg	ccttgtagc	aaagtttttg	ttgaaagctg	tgctcagagg	aggtgagagg	240
agagggaagg	gaaaactgca	tcataacttt	acagaattga	atctagagtc	ttccccgaaa	300
agccccagaa	cttctctgcn	gnatctggct	tgtccatctg	gtctaagggt	gctgcttctt	360
ccccagccat	cgagtcagtt	tgtgcccatc	aataatacac	gaacctgctat	ttcccatgac	424
tgct						
<210>	158					
<211>	2099					
<212>	DNA					
<213>	Homo sapiens					
<400>	158					
cgcgcgttaa	aaggcgcagc	aggtggggagc	cggggccttc	accgcgaacc	cgacgagagc	60
cgcacagccg	gcggcgcccg	agcccgacct	gcctgcccag	ccggagcgaa	ggggcccgcc	120
ccgcgcagag	ccgcgcgcag	ggccgcgcgc	cgacagcgag	ttaaaacgtg	caggcaccag	180
aaggcacttc	ctgtcggtga	agaagacctg	tctccggtgt	cacgggcatc	ctgtgttttg	240
caaacggggc	tgacctccct	tcctggggag	caggaagggt	cagggaagga	aaagaagtac	300
agaaagatctg	gctaaacaat	ttctgtatgg	cgaaagaaaa	attctaactt	gtacgccctc	360
ttcatgcctc	tttaattcaa	tttgaatatt	ccaggcgaca	tcctcactga	ccgagcaaa	420
attgacattc	gtatcatcac	tgtgcaacct	tggtctctag	gcactccagt	ggggtagagg	480
aaggaggtct	gaaacctctg	cagaggggac	ttgcctcat	tccttgggtc	tgaaacactg	540
cgagtcgttg	gaaacaggac	tcagggaata	accagcgaca	tggattgggg	gacgcgtcac	600
actttcatcg	ggggtgtcaa	caaacactcc	accagcatcg	ggaagggtgtg	gatcacagtc	660
atcttttatt	tcocagtcct	gatcctcgtg	gtggctgcc	aggaagtgtg	gggtgacgag	720
caagaggact	tcgtctgcga	cacactgcga	ccgggatgca	aaaatgtgtg	ctatgaccac	780
tttttccogt	tgtcccacat	ccggtctgtg	gcctccagc	tgatcttcgt	ctccaccoca	840
gcgtctgctg	tgggccatga	tgtggcctac	tacaggcacg	aaacctcgca	caagttcagg	900
cgaggagaga	agaggaatga	tttcaaaagc	atagaggaca	ttaaaaagca	gaaggttcgg	960
atagagggtg	cgctgtggtg	gacgtacacc	agcagcatct	ttttccgaat	catctttgaa	1020
gcagcccttta	tgtatgtgtt	ttactttcct	tacaatgggt	accacctgcc	ctgggtgtgtg	1080
aaatgtggga	ttgaccctcg	ccccaacctt	gttgactgct	ttatttctag	gccaacacag	1140
aaagacgtgt	ttacactttt	tatgattttt	gcgtctgtga	tttgcatgct	gccttaacgtg	1200
gcagagttgt	gctgaactgt	gctgaagtgt	tgtttttaga	gatcaaaagc	agcaacagag	1260
caaaaaaatc	accccaata	tgccctaaga	gagagtaagc	agaatgaaat	gaatgagctg	1320
atttcagata	gtggtcaaaa	tgcaatcaca	ggtttccaa	ctaaacattt	caaggtaaaa	1380
tgtagctgcg	tcataaggag	acttctgtct	tctccagaag	gcaataccaa	ctgtaaaagt	1440
cccttctgtg	cctgaagagt	ttgttaaatg	ctttcataat	aaatagacac	ttagtgttaac	1500
tttttctagg	atacttgcct	cattcataca	caacgtaatc	aaatctgtgg	tcacatctgtg	1560
aaaacaagag	actgcttgac	aaaggagcat	tgacgtcact	ttgacaggtt	cccttttaagt	1620

```

ggactctctg acaaagtggg tactttctga aaatttatat aactgttgtt gataaggaac 1680
atttatccag gaattgatac gtttatttagg aaaagatatt tttatagcct tggatgtttt 1740
tagttctgac tttgaattta tataaagtat ttttataatg actgggtcttc cttacctgga 1800
aaaacatgac atgttagttt tagaattaca ccacaagtat ctaaaatttg aacttacaaa 1860
gggtctatct tgtaaatatt gttttgcatt gtctgttggc aaatttgtga actgtcatga 1920
tacgcttaag gtggaagtgt ttcattgcac aatatatttt tactgctttc tgaatgtaga 1980
cggaacagtg tgggaagcaga aggccttttt aactcatcgc ttgccaatc attgcaaaaca 2040
actgaaatgt ggatgtgatt gcctcaataa agctcgtccc cattgcttaa aaaaaaaaaa 2099

```

&lt;210&gt; 159

&lt;211&gt; 291

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 159

```

Met Asp Trp Gly Thr Leu His Thr Phe Ile Gly Gly Val Asn Lys His
 1          5          10          15
Ser Thr Ser Ile Gly Lys Val Trp Ile Thr Val Ile Phe Ile Phe Arg
 20          25          30
Val Met Ile Leu Val Val Ala Ala Gln Glu Val Trp Gly Asp Glu Gln
 35          40          45
Glu Asp Phe Val Cys Asn Thr Leu Gln Pro Gly Cys Lys Asn Val Cys
 50          55          60
Tyr Asp His Phe Phe Pro Val Ser His Ile Arg Leu Trp Ala Leu Gln
 65          70          75          80
Leu Ile Phe Val Ser Thr Pro Ala Leu Leu Val Ala Met His Val Ala
 85          90          95
Tyr Tyr Arg His Glu Thr Thr Arg Lys Phe Arg Arg Gly Glu Lys Arg
 100         105         110
Asn Asp Phe Lys Asp Ile Glu Asp Ile Lys Lys Gln Lys Val Arg Ile
 115         120         125
Glu Gly Ser Leu Trp Trp Thr Tyr Thr Ser Ser Ile Phe Phe Arg Ile
 130         135         140
Ile Phe Glu Ala Ala Phe Met Tyr Val Phe Tyr Phe Leu Tyr Asn Gly
 145         150         155         160
Tyr His Leu Pro Trp Val Leu Lys Cys Gly Ile Asp Pro Cys Pro Asn
 165         170         175
Leu Val Asp Cys Phe Ile Ser Arg Pro Thr Glu Lys Thr Val Phe Thr
 180         185         190
Ile Phe Met Ile Ser Ala Ser Val Ile Cys Met Leu Leu Asn Val Ala
 195         200         205
Glu Leu Cys Tyr Leu Leu Leu Lys Val Cys Phe Arg Arg Ser Lys Arg
 210         215         220
Ala Gln Thr Gln Lys Asn His Pro Asn His Ala Leu Lys Glu Ser Lys
 225         230         235         240
Gln Asn Glu Met Asn Glu Leu Ile Ser Asp Ser Gly Gln Asn Ala Ile
 245         250         255
Thr Gly Ser Gln Ala Lys His Phe Lys Val Lys Cys Ser Cys Val Ile
 260         265         270
Arg Arg Leu Leu Ser Ser Pro Glu Gly Asn Thr Asn Leu Lys Val Pro
 275         280         285
Ser Val Ala
 290

```

&lt;210&gt; 160

&lt;211&gt; 3951

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 160

tctgcatcca	tattgaaaac	ctgcacacaat	gtatgacgca	ggctcagtg	gagtgaactg	60
gaggctctct	tacaacatga	cccaaaaggag	cattgcaggt	cttatttgca	acctgaagtt	120
tgtgactctc	ctggttgcc	taagttcaga	actccattc	ctgggagctg	gagtacagct	180
tcaagacaat	gggtataatg	gattgtctcat	tgcatttaat	cttcaggtatc	ctgagaatca	240
gaactcatc	tcaaacattt	aggaatgat	aactggaagt	tcattttacc	tatttaatgc	300
taccaagaga	agagtatttt	tcagaaatat	aaagatttta	atacctgcc	catggaaaag	360
taataataac	agcaaaataa	aacaagaatc	atatgaaaa	gcaaatgtca	tagtgactga	420
ctggatggg	gcacatggag	atgatccata	cacccataca	tacagaggg	gtggaaaaga	480
gggaaaatac	attcatttca	cacotaattt	ctactgaat	gataacttaa	cagctggcta	540
cggatcacga	ggcggagtgt	ttgtccatga	atgggccac	ctccgtggg	gtgtgttcca	600
tgagtataac	aatgacaaac	ctttctacat	aaatgggcaa	aatcaaat	aagtgcacag	660
gtgtcatct	gacatcacag	gcatttttgt	gtgtgaaaa	ggctcctg	cccaagaaaa	720
ctgtattatt	agtaagcttt	ttaaaagaag	atgcaccttt	atctacaata	gcacccaaaa	780
tgcaactgca	tcaataatgt	tcatgcaaa	tttatctct	gtggttgaa	tttgtaatgc	840
aagtaaccac	aaccaagag	caccaaacct	acagaaacag	atgtgcagcc	tcagaagtgc	900
atgggatgta	atccagagct	ctgctgaact	taccacagc	tttcccatga	acgggactga	960
gcttccact	cctccacat	tctcgctgt	agagctggt	gacaaagtgc	gtctttgatt	1020
gctggatgtg	tccagcaaga	tggcagaggc	tgacagactc	cttcaactac	acaagacgcg	1080
agaattttat	ttgatgcaga	tttgtgaaat	ctataccttc	gtgggcattg	ccagtttcga	1140
cagcaaaagg	gagatcagag	ccagctaca	ccaaattaac	agcaatgatg	atcgaagatt	1200
gctggtttca	tatctgccca	ccactgtatc	agctaaaaca	gacatcagca	tttgtcagg	1260
gcttaagaaa	ggatttgagg	tggttgaaaa	actgaatgga	aaagcttatg	gtctgtgat	1320
gatattagt	accagcggag	atgataagct	tcttggaat	tgcttacc	ctgtgctcag	1380
cagtggttca	acaattcatc	ccattgccct	gggtcatct	cgagcccaaa	atctggagga	1440
atatacagt	cttcacaggag	gtttaaagtt	ctttgtcca	gatattatca	gactcaatag	1500
catgattgat	gctttcagta	gaatttcttc	tggaaactgga	gacattttcc	agcaacatat	1560
tcagcttgaa	agtacagggt	aaaatgtcaa	acctcacact	caattgaaaa	acacagtgac	1620
tgtggataat	actgtgggca	acgacactat	gtttctagtt	acgtggcagg	ccagtggtcc	1680
tctgagatt	atattatttg	atcctgatgg	acgaaaatc	tacacaaata	attttatcac	1740
caatctaact	tttcggacag	ctagtcttg	gattccagga	acagctaagc	cttgggcactg	1800
gacttacacc	ctgaacaata	cccatcattc	tctgcaagcc	ctgaaagtga	cagtgacctc	1860
tcgcgcctcc	aactcagctg	tgccccagc	cactgtggaa	gcctttgttg	aaagagagag	1920
cctccatttt	cctcatcctg	tgatgattta	tgccaatgtg	aaacagggat	tttatcccat	1980
tcttaatgcc	actgtcactg	ccacagttga	gccagagact	ggagatcctg	tatcagctgag	2040
actccttgat	gatggagcag	gtgctgatgt	tataaaaaat	gatggaattt	actcgaggta	2100
ttttttctcc	tttgtgcgaa	atggttagata	tagcttgaaa	gtgcagtgtca	atcactctcc	2160
cagatcaagg	accocagccc	actctattcc	agggagtcat	gctatgtatg	taccaggtta	2220
cacagcaaac	ggtaattatc	agatgaatgc	tccaaggaaa	tcagtaggca	gaatagagga	2280
ggagcgaagg	tggggcttta	gccgagtcag	ctcaggagga	ctcctttcag	gtctgggagt	2340
tccagctggc	ccccaccctg	atgtgtttcc	acctgcaaa	attattgacc	tggaaagctg	2400
aaaagttaga	gaggaattga	ccctatcttg	gacagcaact	ggagaagaat	ttgatcaggg	2460
ccaggctaca	agctatgaaa	taagaatgag	taaaagtcta	cagaatatcc	aaagatgact	2520
taacaatgct	attttagtaa	atacatcaaa	gcgaaatcct	cagcaagctg	gcaatcaggga	2580
gatatttcag	ttctcacccc	aaattttocac	gaatggactc	gaacatcagc	caaatggaga	2640
aacacatgaa	agccacagaa	tttatgttgc	aatcacgaca	atggacagga	actccttaca	2700
gtctgctgta	tctaaacttg	cccaggccoc	ctgttttatt	cccccaact	tgatctctgt	2760
acctgcocga	gattatctta	tattgaaagg	agttttaaca	gcaatggggt	tgatagggaat	2820
catttgcctt	attatagttg	tgacacatca	tactttaagc	agggaaaaaa	gagcagacaa	2880
gaaagagcgt	ggaaacaaat	tattataaat	aaatatccaa	agtgctcttc	ttcttagata	2940
taagaccocat	ggcctctcag	tacaaaaaca	tactaacaaa	gtcaaattaa	catcaaaact	3000
gtattataat	gcattgagtt	tttgtacaat	acagataaga	tttttcaatg	gtgatacaac	3060
aaattctttt	tgggggtaga	ttgaaaaacc	cttacacttt	ggctatgaac	aaataataaa	3120
aattattctt	taagaataat	tctttaaagg	caaaaggga	ggtaaaagtc	gaccagtgct	3180
aaggaaagtt	tgttttattg	aggtgaaaaa	atagcccaaa	gcagagaaaa	gggggtagg	3240
tctgcattat	aactgtctgt	gtgaagcaat	catttagtta	ctttgattaa	tttttctttt	3300



```

ctccttatct gtcgagaaca ggttgcttgt ttacaactga agatcatgct atatttcata 3360
tatgaagccc ctaatgcaaa gctctttacc tcttgctatt ttgttatata tattacagat 3420
gaaatctcac tgctaagtct cagagatctt ttttcaactgt aagaggtaac ctttaacaat 3480
atgggtatta cctttgtctc ttcataccgg ttttatgaca aaggtctatt gaatttat 3540
gttttgaagt ttctactccc atcaaagcag ctttttaagt tattgctctg gttattatgg 3600
atgatagtta tagcccttat atgccttaa ctaaggaaga aaagatgtta ttctgagttt 3660
gttttaatac atatatgaac atatagtttt attcaattaa accaaagaag aggtcagcag 3720
ggagatacta acctttggaa atgattagct ggcctctgttt ttgtgttaaa taagagtcct 3780
taactccttc tocatcaaga gttacttacc aagggcaggg gaaggggat atagaggtcc 3840
caaggaaata aaaatcatct ttcactttaa attttactcc ttctcttat ttttttaaaa 3900
gattatcgaa caataaaatc atttgctttt ttaattaaaa acataaaaaa a 3951

```

<210> 161  
 <211> 943  
 <212> PRT  
 <213> Homo sapiens

<400> 161

Met	Thr	Gln	Arg	Ser	Ile	Ala	Gly	Pro	Ile	Cys	Asn	Leu	Lys	Phe	Val
1			5					10					15		
Thr	Leu	Leu	Val	Ala	Leu	Ser	Ser	Glu	Leu	Pro	Phe	Leu	Gly	Ala	Gly
			20					25					30		
Val	Gln	Leu	Gln	Asp	Asn	Gly	Tyr	Asn	Gly	Leu	Leu	Ile	Ala	Ile	Asn
			35				40					45			
Pro	Gln	Val	Pro	Glu	Asn	Gln	Asn	Leu	Ile	Ser	Asn	Ile	Lys	Glu	Met
			50				55				60				
Ile	Thr	Glu	Ala	Ser	Phe	Tyr	Leu	Phe	Asn	Ala	Thr	Lys	Arg	Arg	Val
			65			70			75				80		
Phe	Phe	Arg	Asn	Ile	Lys	Ile	Leu	Ile	Pro	Ala	Thr	Trp	Lys	Ala	Asn
			85					90					95		
Asn	Asn	Ser	Lys	Ile	Lys	Gln	Glu	Ser	Tyr	Glu	Lys	Ala	Asn	Val	Ile
			100					105					110		
Val	Thr	Asp	Trp	Tyr	Gly	Ala	His	Gly	Asp	Asp	Pro	Tyr	Thr	Leu	Gln
			115				120					125			
Tyr	Arg	Gly	Cys	Gly	Lys	Glu	Gly	Lys	Tyr	Ile	His	Phe	Thr	Pro	Asn
			130				135				140				
Phe	Leu	Leu	Asn	Asp	Asn	Leu	Thr	Ala	Gly	Tyr	Gly	Ser	Arg	Gly	Arg
			145			150			155				160		
Val	Phe	Val	His	Glu	Trp	Ala	His	Leu	Arg	Trp	Gly	Val	Phe	Asp	Glu
			165					170					175		
Tyr	Asn	Asn	Asp	Lys	Pro	Phe	Tyr	Ile	Asn	Gly	Gln	Asn	Gln	Ile	Lys
			180				185					190			
Val	Thr	Arg	Cys	Ser	Ser	Asp	Ile	Thr	Gly	Ile	Phe	Val	Cys	Glu	Lys
			195				200					205			
Gly	Pro	Cys	Pro	Gln	Glu	Asn	Cys	Ile	Ile	Ser	Lys	Leu	Phe	Lys	Glu
			210				215				220				
Gly	Cys	Thr	Phe	Ile	Tyr	Asn	Ser	Thr	Gln	Asn	Ala	Thr	Ala	Ser	Ile
			225			230			235				240		
Met	Phe	Met	Gln	Ser	Leu	Ser	Ser	Val	Val	Glu	Phe	Cys	Asn	Ala	Ser
			245						250				255		
Thr	His	Asn	Gln	Glu	Ala	Pro	Asn	Leu	Gln	Asn	Gln	Met	Cys	Ser	Leu
			260				265					270			
Arg	Ser	Ala	Trp	Asp	Val	Ile	Thr	Asp	Ser	Ala	Asp	Phe	His	His	Ser
			275				280					285			
Phe	Pro	Met	Asn	Gly	Thr	Glu	Leu	Pro	Pro	Pro	Pro	Thr	Phe	Ser	Leu
			290				295				300				
Val	Glu	Ala	Gly	Asp	Lys	Val	Val	Cys	Leu	Val	Leu	Asp	Val	Ser	Ser
			305			310				315					320

Lys Met Ala Glu Ala Asp Arg Leu Leu Gln Leu Gln Gln Ala Ala Glu  
 325 330 335  
 Phe Tyr Leu Met Gln Ile Val Glu Ile His Thr Phe Val Gly Ile Ala  
 340 345 350  
 Ser Phe Asp Ser Lys Gly Glu Ile Arg Ala Gln Leu His Gln Ile Asn  
 355 360 365  
 Ser Asn Asp Asp Arg Lys Leu Leu Val Ser Tyr Leu Pro Thr Thr Val  
 370 375 380  
 Ser Ala Lys Thr Asp Ile Ser Ile Cys Ser Gly Leu Lys Lys Gly Phe  
 385 390 395 400  
 Glu Val Val Glu Lys Leu Asn Gly Lys Ala Tyr Gly Ser Val Met Ile  
 405 410 415  
 Leu Val Thr Ser Gly Asp Asp Lys Leu Leu Gly Asn Cys Leu Pro Thr  
 420 425 430  
 Val Leu Ser Ser Gly Ser Thr Ile His Ser Ile Ala Leu Gly Ser Ser  
 435 440 445  
 Ala Ala Pro Asn Leu Glu Glu Leu Ser Arg Leu Thr Gly Gly Leu Lys  
 450 455 460  
 Phe Phe Val Pro Asp Ile Ser Asn Ser Asn Ser Met Ile Asp Ala Phe  
 465 470 475 480  
 Ser Arg Ile Ser Ser Gly Thr Gly Asp Ile Phe Gln Gln His Ile Gln  
 485 490 495  
 Leu Glu Ser Thr Gly Glu Asn Val Lys Pro His His Gln Leu Lys Asn  
 500 505 510  
 Thr Val Thr Val Asp Asn Thr Val Gly Asn Asp Thr Met Phe Leu Val  
 515 520 525  
 Thr Trp Gln Ala Ser Gly Pro Pro Glu Ile Ile Leu Phe Asp Pro Asp  
 530 535 540  
 Gly Arg Lys Tyr Tyr Thr Asn Asn Phe Ile Thr Asn Leu Thr Phe Arg  
 545 550 555 560  
 Thr Ala Ser Leu Trp Ile Pro Gly Thr Ala Lys Pro Gly His Trp Thr  
 565 570 575  
 Tyr Thr Leu Asn Asn Thr His His Ser Leu Gln Ala Leu Lys Val Thr  
 580 585 590  
 Val Thr Ser Arg Ala Ser Asn Ser Ala Val Pro Pro Ala Thr Val Glu  
 595 600 605  
 Ala Phe Val Glu Arg Asp Ser Leu His Phe Pro His Pro Val Met Ile  
 610 615 620  
 Tyr Ala Asn Val Lys Gln Gly Phe Tyr Pro Ile Leu Asn Ala Thr Val  
 625 630 635 640  
 Thr Ala Thr Val Glu Pro Glu Thr Gly Asp Pro Val Thr Leu Arg Leu  
 645 650 655  
 Leu Asp Asp Gly Ala Gly Ala Asp Val Ile Lys Asn Asp Gly Ile Tyr  
 660 665 670  
 Ser Arg Tyr Phe Phe Ser Phe Ala Ala Asn Gly Arg Tyr Ser Leu Lys  
 675 680 685  
 Val His Val Asn His Ser Pro Ser Ile Ser Thr Pro Ala His Ser Ile  
 690 695 700  
 Pro Gly Ser His Ala Met Tyr Val Pro Gly Tyr Thr Ala Asn Gly Asn  
 705 710 715 720  
 Ile Gln Met Asn Ala Pro Arg Lys Ser Val Gly Arg Asn Glu Glu Glu  
 725 730 735  
 Arg Lys Trp Gly Phe Ser Arg Val Ser Ser Gly Gly Ser Phe Ser Val  
 740 745 750  
 Leu Gly Val Pro Ala Gly Pro His Pro Asp Val Phe Pro Pro Cys Lys  
 755 760 765  
 Ile Ile Asp Leu Glu Ala Val Lys Val Glu Glu Glu Leu Thr Leu Ser  
 770 775 780

Trp	Thr	Ala	Pro	Gly	Glu	Asp	Phe	Asp	Gln	Gly	Gln	Ala	Thr	Ser	Tyr	
785					790					795					800	
Glu	Ile	Arg	Met	Ser	Lys	Ser	Leu	Gln	Asn	Ile	Gln	Asp	Asp	Phe	Asn	
			805					810						815		
Asn	Ala	Ile	Leu	Val	Asn	Thr	Ser	Lys	Arg	Asn	Pro	Gln	Gln	Ala	Gly	
			820					825					830			
Ile	Arg	Glu	Ile	Phe	Thr	Phe	Ser	Pro	Gln	Ile	Ser	Thr	Asn	Gly	Pro	
			835				840					845				
Glu	His	Gln	Pro	Asn	Gly	Glu	Thr	His	Glu	Ser	His	Arg	Ile	Tyr	Val	
			850			855					860					
Ala	Ile	Arg	Ala	Met	Asp	Arg	Asn	Ser	Leu	Gln	Ser	Ala	Val	Ser	Asn	
			865		870					875				880		
Ile	Ala	Gln	Ala	Pro	Leu	Phe	Ile	Pro	Pro	Asn	Ser	Asp	Pro	Val	Pro	
			885					890						895		
Ala	Arg	Asp	Tyr	Leu	Ile	Leu	Lys	Gly	Val	Leu	Thr	Ala	Met	Gly	Leu	
			900					905					910			
Ile	Gly	Ile	Ile	Cys	Leu	Ile	Ile	Val	Val	Thr	His	His	Thr	Leu	Ser	
			915			920						925				
Arg	Lys	Lys	Arg	Ala	Asp	Lys	Lys	Glu	Asn	Gly	Thr	Lys	Leu	Leu		
	930					935					940					

&lt;210&gt; 162

&lt;211&gt; 498

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 162

tggagaacca	cgtggacagc	accatgaaca	tgttggcgcg	gggaggcagt	gctggccgga	60
agccctc0aa	gtcgggtatg	aaggagctgg	ccgtgttcog	ggagaaggtc	actgagcagc	120
accggcagat	gggcaaggtt	ggcaagcatc	accttggcct	ggaggagccc	aagaagctgc	180
gaccacc0cc	tgc0aggact	ccotgccaac	aggaactgga	ccaggtcctg	gagcggatct	240
ccaccatcgc	ccctccggat	gagcggggcc	ctctggagca	ccctactcc	ctgcacatcc	300
ccaactgtga	caagcatggc	ctgtacaacc	tcaaacagtg	gcaagatgic	tctgaacggg	360
cagcgtgggg	agtgtcgggt	tgtgaacccc	aacaccggga	agctgatcca	gggagccccc	420
accatccggg	gggacc0cga	gtgtcatctc	ttctacaatg	agcagcagga	ggctcgcggg	480
gtgcacaccc	cagcggat					498

&lt;210&gt; 163

&lt;211&gt; 1128

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 163

gccac0ctgg	cctcctgato	gacgacacac	gcacttgaaa	cttgcttctca	gggtgtgtgg	60
aatacaactt	ccggaagcaa	ccagccc0acc	agaggaggtc	ccgagccgcga	gcggagacga	120
tgcagcggag	actgggttcag	cagtgagcgc	tcgcgggtgt	cctgtcagc	ta0cgggtgc	180
cctcctcggg	gcgctcgggt	gagggtctca	gcc0ccgcct	caaaagagct	gtgtctgaac	240
atcagctcct	ccatgacaag	gggaagtcca	tccaagattt	acggcgacga	ttctccttc	300
accatctgat	cgcagaaatc	cacacagctg	aaatcagagc	tacctcggag	gtgtcc0cta	360
actccaagcc	ctctcc0aac	acaaagaacc	accc0ctc0g	atttgggtct	gatgatgagg	420
gcagatcact	aactcaggaa	actaacaagg	tggagacgta	caaagagcag	c0gctcaaga	480
caactgggaa	gaaaaagaaa	ggcaagcc0g	ggaacgc0aa	ggagcaggaa	aagaaaaaac	540
ggcgaactcg	ctctgcctgg	ttagactctg	gagtgactgg	gagtgggcta	gaaggggacc	600
acctgtctga	cac0tccaca	acgtc0ctgg	agctc0gattc	acggaggcat	tgaatttttc	660
agcagagacc	ttccaaggac	atattgcagg	attctgtaat	agtgaacata	tggaaagatg	720
tagaataatt	tattgtcctg	aaatactgtg	aatgcatttg	aataaaactg	tctcc0cat	780
tgctctatga	aactgcacat	tggtcattgt	gaatattttt	ttttttgcc	aggctaatacc	840

```

aattattatt atcacattta ccataattta tttgtccat tgatgtattt attttgtaa 900
tgtatcttgg tgctgctgaa tttctatatt ttttgttaaca taatgcactt tagatataca 960
tatcaagatg gttgataaat gacacaatga agtgtctcta tttgtgggtt gattttaatg 1020
aatgcctaaa tataattatc caaattgatt ttcccttgtg catgtaaaaa taacagtatt 1080
ttaaatttgg aaagaatgtc taataaaata taatctaatt acatcatg 1128

```

&lt;210&gt; 164

&lt;211&gt; 1310

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 164

```

gggcctgggt cgcaagaag ctgacttcag agggggaaac tttcttcttt taggaggcgg 60
ttagccctgt tocacgaacc caggagaact gctggccaga ttaattagac attgctatgg 120
gagacgtgta aacacactac ttatcattga tgcataataa aaacattttt attttcgcta 180
ttatttcaga ggaagcgctt ctgatttgggt tctttttttc ctttttgtct tttctggctg 240
tgtgttttgg agaagacaca gttggagtga ccggttgcta aataagctcc gagcgcgagc 300
ggagacgatg cagcggagac tggttcagca gtggagcgtc gcggtgttcc tgcgtagcta 360
cgcggtgccc tctgcggggc gctcgttgga ggtctcagc cgcgcctca aaagagctgt 420
gtctgaacat cagctcctcc atgacaaggg gaagtccat caagatttac ggcgcagcatt 480
cttctctcac catctgatcg cagaaatcca cacagctgaa atcagagcta cctcggaggt 540
gtccctcaac tocaagccct ctcccaacac aaagaaccac ccgctccgat ttgggtctga 600
tgatgagggc agatacctaa ctcaggaaac taacaaggtg gagacgtaca aagagcagcc 660
gctcaagaca cctgggaaga aaaagaaggc caagcccggg aaacgcaagg agcaggaaaa 720
gaaaaaacgg gaaactcgtc ctgcctgggt agactctgga gtgactggga gtgggctaga 780
aggggaccac ctgtctgaca cctccacaac gtgcgtggag ctcgattcac ggaggcattg 840
aaattttcag cagagacctt ocaaggacat attgcaggat tctgtaatag tgaacatatg 900
gaaagtatta gaaatatatta ttgtctgtaa atactgtaa tgcattggaa taaaactgtc 960
tccccattgt ctctatgaaa ctgcacattg gtcatgtga atattttttt ttttgccaa 1020
gctaatacaa ttattattat cacatttacc ataatttatt ttgtcattg atgtatttat 1080
tttgtaaatg tatcttgggt ctgcgtaatt tctatatttt ttgtaacata atgcacttta 1140
gatatacata tcaagtatgt tgataaatga cacaatgaag tgcctctatt ttgtggttga 1200
ttttaatgaa tgctaaata taattatcca aattgtattt cctttgtgccc cgtaaaaaata 1260
acagtatttt aaatttggaa agaattgtca ataaaaata atctaattac 1310

```

&lt;210&gt; 165

&lt;211&gt; 177

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 165

```

Met Gln Arg Arg Leu Val Gln Gln Trp Ser Val Ala Val Phe Leu Leu
  1          5          10          15
Ser Tyr Ala Val Pro Ser Cys Gly Arg Ser Val Glu Gly Leu Ser Arg
  20          25          30
Arg Leu Lys Arg Ala Val Ser Glu His Gln Leu Leu His Asp Lys Gly
  35          40          45
Lys Ser Ile Gln Asp Leu Arg Arg Phe Phe Leu His His Leu Ile
  50          55          60
Ala Glu Ile His Thr Ala Glu Ile Arg Ala Thr Ser Glu Val Ser Pro
  65          70          75          80
Asn Ser Lys Pro Ser Pro Asn Thr Lys Asn His Pro Val Arg Phe Gly
  85          90          95
Ser Asp Asp Glu Gly Arg Tyr Leu Thr Gln Glu Thr Asn Lys Val Glu
 100          105          110
Thr Tyr Lys Glu Gln Pro Leu Lys Thr Pro Gly Lys Lys Lys Gly
 115          120          125
Lys Pro Gly Lys Arg Lys Glu Gln Glu Lys Lys Lys Arg Thr Arg

```

130		135		140
Ser Ala Trp Leu Asp	Ser Gly Val Thr Gly	Ser Gly Leu Glu Gly Asp		
145	150	155	160	
His Leu Ser Asp Thr	Ser Thr Thr Ser Leu Glu	Leu Asp Ser Arg Arg		
	165	170	175	

His

<210> 166  
 <211> 177  
 <212> PRT  
 <213> Homo sapiens

<400> 166

Met Gln Arg Arg Leu Val Gln Gln Trp Ser Val Ala Val Phe Leu Leu	
1	5 10 15
Ser Tyr Ala Val Pro Ser Cys Gly Arg Ser Val Glu Gly Leu Ser Arg	
	20 25 30
Arg Leu Lys Arg Ala Val Ser Glu His Gln Leu Leu His Asp Lys Gly	
	35 40 45
Lys Ser Ile Gln Asp Leu Arg Arg Arg Phe Phe Leu His His Leu Ile	
	50 55 60
Ala Glu Ile His Thr Ala Glu Ile Arg Ala Thr Ser Glu Val Ser Pro	
65	70 75 80
Asn Ser Lys Pro Ser Pro Asn Thr Lys Asn His Pro Val Arg Phe Gly	
	85 90 95
Ser Asp Asp Glu Gly Arg Tyr Leu Thr Gln Glu Thr Asn Lys Val Glu	
	100 105 110
Thr Tyr Lys Glu Gln Pro Leu Lys Thr Pro Gly Lys Lys Lys Gly	
	115 120 125
Lys Pro Gly Lys Arg Lys Glu Gln Glu Lys Lys Lys Arg Arg Thr Arg	
	130 135 140
Ser Ala Trp Leu Asp Ser Gly Val Thr Gly Ser Gly Leu Glu Gly Asp	
145	150 155 160
His Leu Ser Asp Thr Ser Thr Thr Ser Leu Glu Leu Asp Ser Arg Arg	
	165 170 175

His

<210> 167  
 <211> 3362  
 <212> DNA  
 <213> Homo sapiens

<400> 167

cacaatgtat gcagcaggct cagtgtgagt gaactggagg cttctctaca acatgaccca	60
aaggagcatt gcaggctcta ttigcaacct gaagtttgtg actctctctg ttgccttaag	120
ttcagaactc ccatttcttg gagctggagt acagcttcaa gacaatgggt ataattggatt	180
gtctattgca attaatcttc aggtacctga gaatcagaac ctcatctcaa acattaaagga	240
aatgataact gaagcttcat tttaacctatt taatgctaacc aagagaagag tatttttcag	300
aaatataaag attttaatac ctgccacatg gaaagctaat aataacagca aataaaaaa	360
agaatcatat gaaaaggcaa atgtcatagt gactgactgg tatggggcac atggagatga	420
tcatacacacc ctacaataca gaggggtgtgg aaaagaggga aaatacattc atttcacacc	480
taatttctta ctgaatgata acttaacagc tggctacgga tcacgaggcc gagtgtttgt	540
coatgaatgg gccacacctc gttgggggtg gttcgatgag tataacaatg acaaaccttt	600
ctacataaat gggcaaaatc aaattaaagt gacaagggtg tcactctgaca tcacaggcat	660

ttttgtgtgt	gaaaaaggtc	cttgccccc	agaaaaactgt	attattagta	agctttttaa	720
agaaggatgc	acottttatct	acaatagcac	ccaaaaatgca	actgcatcaa	taatgttcat	780
gcaaagttta	tottctgtggt	ttgaattttg	taatgcaagt	accacacacc	aagaagcacc	840
aaacotacag	aacccagatgt	gcagccctcag	aagtgcattg	gatgtaatca	cagactctgc	900
tgactttcac	cacagcttttc	ccatgaacgg	gactgagctt	ccacctctctc	ccacattctc	960
gctttgtagag	gctgggtgaca	aagtggctctg	tttagtgctg	gatgtgtcca	gcgaagtatg	1020
agaggctgc	agactctcttc	aactacaaca	agccgcagaa	ttttatttga	tgcaagtattg	1080
tgaaattcat	acottctgtgg	gcattgccag	ttctgcacgc	aaaggagaga	tcagagccca	1140
gctacaccaa	attaacagca	atgatgatcg	aaagttgctg	gtttcatatc	tgcccaccac	1200
tgatcacagt	aaaacagaca	tcagcatttg	ttcagggtct	aagaaaggat	ttgaggtggg	1260
tgaaaaactg	aatggaaaag	cttatggctc	tgtgatgata	ttagtgcacca	gcggagatga	1320
taagcttctt	ggcaattgct	taccocactgt	gctcagcagt	ggttcaacaa	ttcactccat	1380
tgccctgggt	tcactctcgag	ccccaaatct	ggaggaatta	tcactgttta	caggaggttt	1440
aaagtctttt	gttccagata	tatcaaaact	caatagcatg	attgatgctt	tcagtagaat	1500
ttctctctgga	actggagaca	ttttccagca	acatattcag	cttgaaagta	caggtgaaaa	1560
tgtaaaactg	caccatcaat	tgaaaaaac	agtgcactgt	gataatactg	tgggcaacga	1620
caactatgtt	ctagttaactg	ggcaggccag	tggtctctct	gagattatat	tatttgatcc	1680
tgatgagcga	aaatactaca	caaataattt	tatcaccaat	ctaacttttc	ggacagctag	1740
cttttggatt	ccagggaacg	ctaagcctgg	gcactggact	tacacccctga	tgtgtttcca	1800
ccatgcocaa	ttattgacct	ggaagctgta	aaagttagaa	aggaattgac	cctatcttgg	1860
acagcacctg	gagaagacct	tgatcagggc	caggctacaa	gctatgaaat	aaagatgagt	1920
aaaaagtctac	agaatatcca	agatgacttt	aacaatgcta	ttttagtaaa	tacatcaaa	1980
cgaaatctct	agcaagctgg	catcaggggag	atattttaact	ttctacccca	aaattccacg	2040
aatggacctg	aacatcagcc	aaatggagaa	acacatgaaa	gccacagaat	tattgttgca	2100
atacagacaa	tgatagtgaa	ctccttacag	tctgctgtat	ctaacattgc	ccaggccctt	2160
ctgtttatct	cccccaattc	tgatcctgta	cctgccagag	attatcttat	attgaaagga	2220
gtttttaacag	caatggggtt	gataggaatc	atttgcttta	ttatagttgt	gatacaatcat	2280
acttttaagca	ggaaaaagag	agcagacaag	aaagagaatg	gaacaaaatt	ttatataata	2340
aaatatccaaa	gtgtctctct	tcttagatat	aagaccocat	gctctcgact	acaaaaacat	2400
actaacacaa	ctcaaatcaac	atcaaaaactg	tattaaaatg	ccattgagtt	ttgtacaata	2460
cagataaagt	ttttacatgg	tagatcaaca	aattcttttt	gggggtagat	tagaaaaacc	2520
ttacactttg	gctatgaaca	aataataaaa	attattcttt	aaagtaatgt	ctttaaaggc	2580
aaagggaaag	gtaaaagtgg	accagtgta	aggaagattt	gttttatgat	gggtgaaaaa	2640
tagccccaag	cagagaaaag	gagggttagt	ctgcattata	actgtctgtg	tgaagcaatc	2700
attttatgat	tttgattaat	ttttcttttc	tcctttctgt	tcgacagacg	gttgctgttt	2760
tacaactgaa	gatcatgcta	tatttcatat	atgaagcccc	taatgcacag	ctctttactt	2820
cttggctat	ttgtatatat	attacagatg	aaatctcaat	gctaagtctc	agagatcttt	2880
tttactgtga	agaggtaacc	tttaacaata	tggttattac	ctttgtctct	tcataccggg	2940
tttatgacaa	aggtctatgt	aatttatgtt	tttgtaagtt	tcactctcca	tcaaggcagc	3000
ttttctgatt	attgcctgtg	ttattatgga	tgatagttat	agcccttata	atgccttaac	3060
taaggagaaga	aagatgttat	tctgagtttg	ttttaataca	tatatgaaca	tatagtttta	3120
ttcaattaaa	ccaaagaaga	ggtcagcagg	gagatactaa	ccottggaaa	tgattagctg	3180
gctctgtttt	ttggttaaat	aagagctctt	aatcctttct	ccatcaagag	ttacttacca	3240
agggcagggt	aagggggata	tagaggtcac	aaggaaataa	aaatcatctt	tcattctttaa	3300
ttttactctg	ttctcttatt	tttttaaaag	attatcgaac	aatcaaatca	tttgcttttt	3360
tt						3362

&lt;210&gt; 168

&lt;211&gt; 2784

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 168

tttgcatcca	tattgaaaac	ctgacacaat	gtatgcagca	ggctcagttg	gagtgaaactg	60
gaggctctct	tacaacatga	cccaaaggag	cattgcaggt	cttattttgca	acctgaagtt	120
tgtgactctc	ctgggttgct	taagttccga	actcccattc	ctggggagctg	gagtcacagt	180
tccaagcaat	gggtataaag	gattgtctat	tgaacattaa	cttcagggtac	tgatgaatca	240
gaacctcatc	tcaaacatta	aggaatgat	aactgaagct	tcattttacc	tattttaatgc	300

```

taccacagaga agagtatttt tcagaaatat aaagatttta atacctgcca catggaaaagc 360
taataataaac agcaaaaataa aacagaatac atagaaaaag gcaaatgtca tagtgactga 420
ctggatgggg gcacatggag atgatccata caccctacaa tacagagggg gtggaaaaaga 480
gggaaaatac attcatttca cactaaattt cctactgaat gataaactta cagctgggcta 540
cggatcacga ggccgagtggt ttgtccatga atggggccac ctcggttggg gtgtgttcga 600
tgagtataac aatgacaaac ctttctacat aaatggggcaa aatcaaatla aagtgaacaag 660
gtgttcatct gacatccacag gccatttttgt gtgtgaaaaa ggctccttgc cccaagaaaa 720
ctgtattatt agtaagcttt ttaagaaggt atgcaccttt atctacaata gcaacccaaaa 780
tgcaactgca tcaataatgt tcatgcaaaag tttatottct gtgggtgaat ttgttaatgc 840
aagtaaccoc aaccaagaag caccaaaacct acagaaccag atgtgcagcc tcagaagtgc 900
atgggatgta atcacagact ctgctgaact tcaccacagc ttcccatga acgggactga 960
gottccacot cctccacat tctcgttgtt agaggctggt gacaaagtgg tctgtttagt 1020
gctggatgtg tccagcaaga tggcagaggt tgacagactc cttcaactac aacaagccgc 1080
agaattttat ttgatgcaga ttgttgaat tcataccttc gtgggcattg ccagtttoga 1140
cagcaaaagga gagatcagag cccagctaca ccaaatlaac agcaaatgat atcgaaagt 1200
gctggtttca tatctgcca ccaactgtatc agctaaaaa gacatcagca ttgtttcag 1260
gottaagaaa ggattttgagg ttggttgaata actgaatgga aaagcttatg gctctgtgat 1320
gatattagtg accagoggag atgataagct tcttggcaat tgcttaccoc actgctgcag 1380
cagtggttca acaattcact ccaattgcctt ggttctcatc tctgttcca gatatatcaa actccaatag 1500
attatcacgt cttacaggag gtttaaaagt cttgttcca gacattttcc agcaacatat 1560
catgattgat gctttcagta gaattttctc ttggaactgga gacattttcc agcaacatat 1560
tcagottgaa agtacaggtg aaaaatgtcaa acctcaccat caattgaaaa acacagtgac 1620
tgtgtgaact acgtgtggga acgacactat gttttcagtt cagtggtgac cagtggtcc 1680
tcttgagatt atattatttg atcctgatgg acgaaaaaac tacacaaata attttatcac 1740
caatctaact tttcggacag ctagtctttg gattccagga acagctaaagc ctgggcactg 1800
gacttacacc ctgaacaata cccatcattc tctgcaagcc ctgaaagtga cagtgacctc 1860
tcggcgctcc aactcagctg tgcocccagc cactgtggaa gcttttggg aaagagacag 1920
cctccatttt cctcatcctg tgatgattta tgccaatgtg aacacagggat ttatcccat 1980
tottaatgcc actgtcactg ccaacagttga gccagagact ggagatcctg ttaacgtgag 2040
actcctgat gatggagcag gtgctgatgt tataaaaaat gatggaattt actogaggtga 2100
ttttttctcc tttgctgcaa atggtagata tagcttgaaa gtgatgtca atcaactctcc 2160
cagcataaac accccagccc actctattcc agggagtcac gctatgtatg taccaggtta 2220
cacagcaaac ggtaatatcc agatgaatgc tccaaggaaa tcagtaggca gaattgagga 2280
ggagcgaaaag tggggcttta gccagagtcag ctacaggagc tccttttcag tgctgggagt 2340
tccagctggc cccacacctg atgtgtttcc accatgcaaa attattgacc tggaaagctgt 2400
aaatagaaga ggaattgacc ctatcttgga cagcacctgg agaagacttt gatcaggggc 2460
aggctacaag ctatgaaata agaattgagta aagatctaca gaatatccaa gatgacttta 2520
acaatgctat tttagtaaat acatcaaagc gaaatcctca gcaagctggc atcagggaga 2580
tatttaagtt ctaaccccaa atttccacga atggacctga acatcagcca aatggagaaa 2640
cacatgaagc ccaagataat tatgttgcaa tacagagtaa ctgttaccag ttcttaccag 2700
ctgctgtatc taacattgcc caggcgctcc tgttttattcc cccaactctc gatcctgtac 2760
ctgccagaga ttatcttata ttga 2784

```

&lt;210&gt; 169

&lt;211&gt; 592

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 169

```

Met Thr Gln Arg Ser Ile Ala Gly Pro Ile Cys Asn Leu Lys Phe Val
1 5 10 15
Thr Leu Leu Val Ala Leu Ser Ser Glu Leu Pro Phe Leu Gly Ala Gly
20 25 30
Val Gln Leu Gln Asp Asn Gly Tyr Asn Gly Leu Leu Ile Ala Ile Asn
35 40 45
Pro Gln Val Pro Glu Asn Gln Asn Leu Ile Ser Asn Ile Lys Glu Met
50 55 60
Ile Thr Glu Ala Ser Phe Tyr Leu Phe Asn Ala Thr Lys Arg Arg Val

```

65					70					75				80
Phe	Phe	Arg	Asn	Ile	Lys	Ile	Leu	Ile	Pro	Ala	Thr	Trp	Lys	Ala
			85						90					95
Asn	Asn	Ser	Lys	Ile	Lys	Gln	Glu	Ser	Tyr	Glu	Lys	Ala	Asn	Val
			100						105					110
Val	Thr	Asp	Trp	Tyr	Gly	Ala	His	Gly	Asp	Asp	Pro	Tyr	Thr	Leu
			115						120					125
Tyr	Arg	Gly	Cys	Gly	Lys	Glu	Gly	Lys	Tyr	Ile	His	Phe	Thr	Pro
			130						135					140
Phe	Leu	Leu	Asn	Asp	Asn	Leu	Thr	Ala	Gly	Tyr	Gly	Ser	Arg	Gly
			145						150					155
Val	Phe	Val	His	Glu	Trp	Ala	His	Leu	Arg	Trp	Gly	Val	Phe	Asp
			165						170					175
Tyr	Asn	Asn	Asp	Lys	Pro	Phe	Tyr	Ile	Asn	Gly	Gln	Asn	Gln	Ile
			180						185					190
Val	Thr	Arg	Cys	Ser	Ser	Asp	Ile	Thr	Gly	Ile	Phe	Val	Cys	Glu
			195						200					205
Gly	Pro	Cys	Pro	Gln	Glu	Asn	Cys	Ile	Ile	Ser	Lys	Lys	Phe	Lys
			210						215					220
Gly	Cys	Thr	Phe	Ile	Tyr	Asn	Ser	Thr	Gln	Asn	Ala	Thr	Ala	Ser
			225						230					235
Met	Phe	Met	Gln	Ser	Leu	Ser	Ser	Val	Val	Glu	Phe	Cys	Asn	Ala
			245						250					255
Thr	His	Asn	Gln	Glu	Ala	Pro	Asn	Leu	Gln	Asn	Gln	Met	Cys	Ser
			260						265					270
Arg	Ser	Ala	Trp	Asp	Val	Ile	Thr	Asp	Ser	Ala	Asp	Phe	His	His
			275						280					285
Phe	Pro	Met	Asn	Gly	Thr	Glu	Leu	Pro	Pro	Pro	Pro	Thr	Phe	Ser
			290						295					300
Val	Glu	Ala	Gly	Asp	Lys	Val	Val	Cys	Leu	Val	Leu	Asp	Val	Ser
			305						310					315
Lys	Met	Ala	Glu	Ala	Asp	Arg	Leu	Leu	Gln	Leu	Gln	Gln	Ala	Ala
			325						330					335
Phe	Tyr	Leu	Met	Gln	Ile	Val	Glu	Ile	His	Thr	Phe	Val	Gly	Ile
			340						345					350
Ser	Phe	Asp	Ser	Lys	Gly	Glu	Ile	Arg	Ala	Gln	Leu	His	Gln	Ile
			355						360					365
Ser	Asn	Asp	Asp	Arg	Lys	Leu	Leu	Val	Ser	Tyr	Leu	Pro	Thr	Thr
			370						375					380
Ser	Ala	Lys	Thr	Asp	Ile	Ser	Ile	Cys	Ser	Gly	Leu	Lys	Lys	Gly
			385						390					395
Glu	Val	Val	Glu	Lys	Leu	Asn	Gly	Lys	Ala	Tyr	Gly	Ser	Val	Met
			405						410					415
Leu	Val	Thr	Ser	Gly	Asp	Asp	Lys	Leu	Leu	Gly	Asn	Cys	Leu	Pro
			420						425					430
Val	Leu	Ser	Ser	Gly	Ser	Thr	Ile	His	Ser	Ile	Ala	Leu	Gly	Ser
			435						440					445
Ala	Ala	Pro	Asn	Leu	Glu	Glu	Leu	Ser	Arg	Leu	Thr	Gly	Gly	Leu
			450						455					460
Phe	Phe	Val	Pro	Asp	Ile	Ser	Asn	Ser	Asn	Ser	Met	Ile	Asp	Ala
			465						470					475
Ser	Arg	Ile	Ser	Ser	Gly	Thr	Gly	Asp	Ile	Phe	Gln	Gln	His	Ile
			485						490					495
Leu	Glu	Ser	Thr	Gly	Glu	Asn	Val	Lys	Pro	His	His	Gln	Leu	Lys
			500						505					510
Thr	Val	Thr	Val	Asp	Asn	Thr	Val	Gly	Asn	Asp	Thr	Met	Phe	Leu
			515						520					525
Thr	Trp	Gln	Ala	Ser	Gly	Pro	Pro	Glu	Ile	Ile	Leu	Phe	Asp	Pro



530						535						540					
Gly	Arg	Lys	Tyr	Tyr	Thr	Asn	Asn	Phe	Ile	Thr	Asn	Leu	Thr	Phe	Arg		
545					550					555					560		
Thr	Ala	Ser	Leu	Trp	Ile	Pro	Gly	Thr	Ala	Lys	Pro	Gly	His	Trp	Thr		
				565					570					575			
Tyr	Thr	Leu	Met	Cys	Phe	His	His	Ala	Lys	Leu	Leu	Thr	Trp	Lys	Leu		
			580					585					590				

&lt;210&gt; 170

&lt;211&gt; 791

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 170

Met	Thr	Gln	Arg	Ser	Ile	Ala	Gly	Pro	Ile	Cys	Asn	Leu	Lys	Phe	Val		
1				5					10					15			
Thr	Leu	Leu	Val	Ala	Leu	Ser	Ser	Glu	Leu	Pro	Phe	Leu	Gly	Ala	Gly		
			20					25					30				
Val	Gln	Leu	Gln	Asp	Asn	Gly	Tyr	Asn	Gly	Leu	Leu	Ile	Ala	Ile	Asn		
		35				40						45					
Pro	Gln	Val	Pro	Glu	Asn	Gln	Asn	Leu	Ile	Ser	Asn	Ile	Lys	Glu	Met		
		50				55						60					
Ile	Thr	Glu	Ala	Ser	Phe	Tyr	Leu	Phe	Asn	Ala	Thr	Lys	Arg	Arg	Val		
65					70					75					80		
Phe	Phe	Arg	Asn	Ile	Lys	Ile	Leu	Ile	Pro	Ala	Thr	Trp	Lys	Ala	Asn		
			85						90					95			
Asn	Asn	Ser	Lys	Ile	Lys	Gln	Glu	Ser	Tyr	Glu	Lys	Ala	Asn	Val	Ile		
			100					105					110				
Val	Thr	Asp	Trp	Tyr	Gly	Ala	His	Gly	Asp	Asp	Pro	Tyr	Thr	Leu	Gln		
		115				120						125					
Tyr	Arg	Gly	Cys	Gly	Lys	Glu	Gly	Lys	Tyr	Ile	His	Phe	Thr	Pro	Asn		
		130				135					140						
Phe	Leu	Leu	Asn	Asp	Asn	Leu	Thr	Ala	Gly	Tyr	Gly	Ser	Arg	Gly	Arg		
145					150					155					160		
Val	Phe	Val	His	Glu	Trp	Ala	His	Leu	Arg	Trp	Gly	Val	Phe	Asp	Glu		
			165						170					175			
Tyr	Asn	Asn	Asp	Lys	Pro	Phe	Tyr	Ile	Asn	Gly	Gln	Asn	Gln	Ile	Lys		
			180					185						190			
Val	Thr	Arg	Cys	Ser	Ser	Asp	Ile	Thr	Gly	Ile	Phe	Val	Cys	Glu	Lys		
		195				200						205					
Gly	Pro	Cys	Pro	Gln	Glu	Asn	Cys	Ile	Ile	Ser	Lys	Leu	Phe	Lys	Glu		
		210				215					220						
Gly	Cys	Thr	Phe	Ile	Tyr	Asn	Ser	Thr	Gln	Asn	Ala	Thr	Ala	Ser	Ile		
225					230					235					240		
Met	Phe	Met	Gln	Ser	Leu	Ser	Ser	Val	Val	Glu	Phe	Cys	Asn	Ala	Ser		
			245						250					255			
Thr	His	Asn	Gln	Glu	Ala	Pro	Asn	Leu	Gln	Asn	Gln	Met	Cys	Ser	Leu		
		260						265					270				
Arg	Ser	Ala	Trp	Asp	Val	Ile	Thr	Asp	Ser	Ala	Asp	Phe	His	His	Ser		
		275					280					285					
Phe	Pro	Met	Asn	Gly	Thr	Glu	Leu	Pro	Pro	Pro	Pro	Thr	Phe	Ser	Leu		
		290				295					300						
Val	Glu	Ala	Gly	Asp	Lys	Val	Val	Cys	Leu	Val	Leu	Asp	Val	Ser	Ser		
305					310					315					320		
Lys	Met	Ala	Glu	Ala	Asp	Arg	Leu	Leu	Gln	Leu	Gln	Gln	Ala	Ala	Glu		
			325					330					335				
Phe	Tyr	Leu	Met	Gln	Ile	Val	Glu	Ile	His	Thr	Phe	Val	Gly	Ile	Ala		

Ser	Phe	Asp	340	Ser	Lys	Gly	345	Ile	Arg	Ala	Gln	Leu	His	Gln	Ile	Asn
		355					360						365			
Ser	Asn	Asp	Asp	Arg	Lys	Leu	Leu	Val	Ser	Tyr	Leu	Pro	Thr	Thr	Val	
		370				375					380					
Ser	Ala	Lys	Thr	Asp	Ile	Ser	Ile	Cys	Ser	Gly	Leu	Lys	Lys	Gly	Phe	
		385			390					395					400	
Glu	Val	Val	Glu	Lys	Leu	Asn	Gly	Lys	Ala	Tyr	Gly	Ser	Val	Met	Ile	
				405					410						415	
Leu	Val	Thr	Ser	Gly	Asp	Asp	Lys	Leu	Leu	Gly	Asn	Cys	Leu	Pro	Thr	
			420					425					430			
Val	Leu	Ser	Ser	Gly	Ser	Thr	Ile	His	Ser	Ile	Ala	Leu	Gly	Ser	Ser	
		435					440					445				
Ala	Ala	Pro	Asn	Leu	Glu	Glu	Leu	Ser	Arg	Leu	Thr	Gly	Gly	Leu	Lys	
		450				455					460					
Phe	Phe	Val	Pro	Asp	Ile	Ser	Asn	Ser	Asn	Ser	Met	Ile	Asp	Ala	Phe	
		465			470					475					480	
Ser	Arg	Ile	Ser	Ser	Gly	Thr	Gly	Asp	Ile	Phe	Gln	Gln	His	Ile	Gln	
				485					490					495		
Leu	Glu	Ser	Thr	Gly	Glu	Asn	Val	Lys	Pro	His	His	Gln	Leu	Lys	Asn	
			500					505					510			
Thr	Val	Thr	Val	Asp	Asn	Thr	Val	Gly	Asn	Asp	Thr	Met	Phe	Leu	Val	
		515					520					525				
Thr	Trp	Gln	Ala	Ser	Gly	Pro	Pro	Glu	Ile	Ile	Leu	Phe	Asp	Pro	Asp	
		530				535					540					
Gly	Arg	Lys	Tyr	Thr	Thr	Asn	Asn	Phe	Ile	Thr	Asn	Leu	Thr	Phe	Arg	
		545			550					555					560	
Thr	Ala	Ser	Leu	Trp	Ile	Pro	Gly	Thr	Ala	Lys	Pro	Gly	His	Trp	Thr	
				565					570					575		
Tyr	Thr	Leu	Asn	Asn	Thr	His	His	Ser	Leu	Gln	Ala	Leu	Lys	Val	Thr	
		580						585					590			
Val	Thr	Ser	Arg	Ala	Ser	Asn	Ser	Ala	Val	Pro	Pro	Ala	Thr	Val	Glu	
		595					600					605				
Ala	Phe	Val	Glu	Arg	Asp	Ser	Leu	His	Phe	Pro	His	Pro	Val	Met	Ile	
		610				615					620					
Tyr	Ala	Asn	Val	Lys	Gln	Gly	Phe	Tyr	Pro	Ile	Leu	Asn	Ala	Thr	Val	
		625			630					635					640	
Thr	Ala	Thr	Val	Glu	Pro	Glu	Thr	Gly	Asp	Pro	Val	Thr	Leu	Arg	Leu	
				645					650					655		
Leu	Asp	Asp	Gly	Ala	Gly	Ala	Asp	Val	Ile	Lys	Asn	Asp	Gly	Ile	Tyr	
			660					665					670			
Ser	Arg	Tyr	Phe	Phe	Ser	Phe	Ala	Ala	Asn	Gly	Arg	Tyr	Ser	Leu	Lys	
		675					680					685				
Val	His	Val	Asn													

&lt;210&gt; 171

&lt;211&gt; 1491

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 171

```

cctcctgcc  gccaagttaa  gacatgctta  cttccoccttc  accctccttc  atgatgtggg  60
aagagtgtct  caaccacgac  ctagccaacg  cgcgatgaga  gggagtgtgc  cgaggggttc  120
tgagaaggtt  tctctcacat  ctagaagaag  gcgcttaaga  tgtggcagcc  cctcttcttc  180
aagtggctct  tgtcctgttg  ccoctgggagt  tctcaaatgt  ctgcagcagc  ctccaccocag  240
cctgaggatg  acatcaatac  acagagggaag  aagagtccag  aaaagatgag  agaagtatac  300
gactctcctg  ggcgaccocg  agagcttacc  attcctcaga  cttcttcaca  tgggtgtaac  360
agattttgtc  ctaaaagtaa  agctctagag  gccgtcaaat  tggcaataga  agccgggttc  420
caccatattg  attctgcaca  tgtttacaat  aatgaggagc  aggttggtact  ggccatccga  480
agcaagattg  cagatggcag  tgtgaagaga  gaagacatat  tctacacttc  aaagctttgg  540
agcaattccc  atcgaccaga  gttggtccga  ccagccttgg  aaaggtcact  gaaaaatctt  600
caattggact  atgttgacct  ctatcttatt  cattttccag  tgtctgtaaa  gccaggtgag  660
gaagtgtacc  caaaagatga  aaatggaaaa  atactatttg  acacagtggg  tctctgtgcc  720
acatggggag  ccatgggaaa  gtgtaaagat  gcaggattgg  ccaagtccat  cgggggtgct  780
aacttcaacc  acagggctgt  ggagatgac  ctcaacaagc  caggggtcaa  gtacaagcct  840
gtctgcaacc  aggtggaatg  tcatccttac  ttcaaccaga  gaaaactgct  ggatttctgc  900
aagtcaaaaag  acattgtctt  ggttgcttat  agtgcctcgt  gatcccatcg  agaagaacca  960
tgggtggacc  cgaactcccc  ggtgctcttg  gaggaccagc  tccctttgtc  cttggcaaaa  1020
aagcacaaag  gaaccccgag  cctgattgcc  ctgcgtacc  agctgcagcg  tgggggtgtg  1080
gtcctggcca  agagctacaa  tgagcagcgc  atcagacaga  acgtgcaggt  gtttgaattc  1140
cagttgactt  cagaggagat  gaaagccata  gatggcctaa  acagaaatgt  gcgatatttt  1200
acccttgata  tttttgtctg  cccocctaatt  tatccatttt  ctgatgaata  ttaacatgga  1260
gggcatttgc  tgaaggtctg  cagaaggccc  tgcgttggtg  tggtagacac  gaggatggct  1320
ctatgctgtg  gactggacac  atgcctctgt  gttaaatctc  tctctgttgg  cgacttcaagt  1380
aagctacagc  taagcccatc  gcccggaaaa  gaaagacaat  aattttgttt  ttcattttga  1440
aaaaattaaa  tgcctctctc  taaagattct  tcacctaaaa  aaaaaaaaaa  a  1491

```

&lt;210&gt; 172

&lt;211&gt; 364

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 172

```

Met Trp Gln Pro Leu Phe Phe Lys Trp Leu Leu Ser Cys Cys Pro Gly
 1          5          10          15
Ser Ser Gln Ile Ala Ala Ala Ser Thr Gln Pro Glu Asp Asp Ile
 20          25          30
Asn Thr Gln Arg Lys Lys Ser Gln Glu Lys Met Arg Glu Val Thr Asp
 35          40          45
Ser Pro Gly Arg Pro Arg Glu Leu Thr Ile Pro Gln Thr Ser Ser His
 50          55          60
Gly Ala Asn Arg Phe Val Pro Lys Ser Lys Ala Leu Glu Ala Val Lys
 65          70          75          80
Leu Ala Ile Glu Ala Gly Phe His His Ile Asp Ser Ala His Val Tyr
 85          90          95
Asn Asn Glu Glu Gln Val Gly Leu Ala Ile Arg Ser Lys Ile Ala Asp
100          105          110
Gly Ser Val Lys Arg Glu Asp Ile Phe Tyr Thr Ser Lys Leu Trp Ser
115          120          125
Asn Ser His Arg Pro Glu Leu Val Arg Pro Ala Leu Glu Arg Ser Leu
130          135          140
Lys Asn Leu Gln Leu Asp Tyr Val Asp Leu Tyr Leu Ile His Phe Pro

```

145				150						155				160	
Val	Ser	Val	Lys	Pro	Gly	Glu	Glu	Val	Ile	Pro	Lys	Asp	Glu	Asn	Gly
				165					170					175	
Lys	Ile	Leu	Phe	Asp	Thr	Val	Asp	Leu	Cys	Ala	Thr	Trp	Glu	Ala	Met
				180				185					190		
Glu	Lys	Cys	Lys	Asp	Ala	Gly	Leu	Ala	Lys	Ser	Ile	Gly	Val	Ser	Asn
				195			200				205				
Phe	Asn	His	Arg	Leu	Leu	Glu	Met	Ile	Leu	Asn	Lys	Pro	Gly	Leu	Lys
				210			215				220				
Tyr	Lys	Pro	Val	Cys	Asn	Gln	Val	Glu	Cys	His	Pro	Tyr	Phe	Asn	Gln
				225			230				235				
Arg	Lys	Leu	Leu	Asp	Phe	Cys	Lys	Ser	Lys	Asp	Ile	Val	Leu	Val	Ala
				245					250					255	
Tyr	Ser	Ala	Leu	Gly	Ser	His	Arg	Glu	Glu	Pro	Trp	Val	Asp	Pro	Asn
				260				265					270		
Ser	Pro	Val	Leu	Leu	Glu	Asp	Pro	Val	Leu	Cys	Ala	Leu	Ala	Lys	Lys
				275			280						285		
His	Lys	Arg	Thr	Pro	Ala	Leu	Ile	Ala	Leu	Arg	Tyr	Gln	Leu	Gln	Arg
				290			295				300				
Gly	Val	Val	Val	Leu	Ala	Lys	Ser	Tyr	Asn	Glu	Gln	Arg	Ile	Arg	Gln
				305			310			315					
Asn	Val	Gln	Val	Phe	Glu	Phe	Gln	Leu	Thr	Ser	Glu	Glu	Met	Lys	Ala
				325					330					335	
Ile	Asp	Gly	Leu	Asn	Arg	Asn	Val	Arg	Tyr	Leu	Thr	Leu	Asp	Ile	Phe
				340				345					350		
Ala	Gly	Pro	Pro	Asn	Tyr	Pro	Phe	Ser	Asp	Glu	Tyr				
				355			360								

&lt;210&gt; 173

&lt;211&gt; 1988

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 173

cgggagcgcg	ctccccgcg	cctcttcgct	tttgtggcgg	cgcccgcgct	cgcaggccac	60
tctctgctgt	gcgccgtccc	gcgcgtcctc	ccgacccgct	ccgctccgct	ccgctcgccg	120
ccgcgcgcgc	cgtcaacatg	atccgctgog	gcctggcctg	ccgagcgctg	ccgtggatcc	180
tgccctcgtc	ctactacagc	gccatcgctc	tgcacatcat	cgcgctggcc	ggccgcggct	240
ggttgccagtc	tagcgccac	ggccagacgt	cctcgctgtg	gtggaaatgc	tcccgaagg	300
cgggcgccag	cggttcctac	gaggaggcgt	gtcagagcct	catggagtag	cggtggggta	360
gagcagcgcc	tgccatgctc	ttctgtggct	tcacatcctc	ggtgatctct	ttcatcctct	420
cctctctcgc	cctctgtgga	ccccagatgc	ttgtcttctc	gagagtgatt	ggaggtctcc	480
ttgccttggc	tgctgtgttc	cagatcatct	ccctggtaat	ttaccocgtg	aagtagaccc	540
agaccttcac	ccttcatgcc	aaccctgctg	tcacttacat	ctataactgg	gcctacggct	600
ttgggtgggc	agccaacgatt	atcctgatcg	gctgtgcctt	cttcttctgc	tgccctccca	660
actacgaaga	tgacctctct	ggcaatgcga	agcccaggta	cttctacaca	ctcgcttaac	720
ttgggaatga	atgtgggaga	aaatcgctgc	tgctgagatg	gactccagaa	gaagaaactg	780
tttctccagg	cgactttgaa	cccatttttt	ggcagtgttc	atattattaa	actagtagaa	840
aatgctaaaa	taatttggga	gaaaatattt	tttaagttag	gttatagttt	catgtttatc	900
ttttattatg	ttttgtgaag	tttgtctctt	tcactaatta	ccataactat	gccaatattt	960
ccttatatct	atccataaca	tttatactac	atttgaaga	gaatatgcac	gtgaaactta	1020
acactttata	aggtaaaaat	gaggtttcca	agatttaata	atctgatcaa	gttcttgtta	1080
tttccaaata	gaatggactt	ggtctgttaa	gggctaagga	gaagaggaag	ataaggttaa	1140
aagttgttaa	tgaccacaaa	ttctaaaaga	aatgcaaaaa	aaaagtttat	tttcaagcct	1200
tcgaactatt	taaggaaaag	aaaatcattt	cctaaaatga	tatcattgtg	gagaattctc	1260
cattaatatc	ctgaatcatt	catttcagct	aaggcttcac	gttgactcga	tatgtcaatc	1320
agggaaatgc	tattttcatg	tccaaaacct	ttgccatagt	tggttaaggc	ttcctttaag	1380

```

tgtgaaatat ttagatgaaa ttttctcttt taaagttctt tatagggtta ggggtgtgga 1440
aaatgctata ttaataaatc tgtagtgttt tgtgtttata tgttcagaac cagagtagac 1500
tggtattgaa gatggactgg gtctaattta tcatgactga tagatctggt taagtgtgtg 1560
agtaaaagcat taggagggtc attcytgtca caaaagtgcc actaaaacag cctcaggaga 1620
ataaatgact tgcttttcta aatctcaggt ttatctgggc tctatcatat agacagggtt 1680
ctgatagttt gcarctgtta gcagaaacct acatatagtt aaaatcctgg tctttcttgg 1740
taaacagatt ttaaattgtct gatataaac atgccacagg agaattcogg accatttagtt 1800
tctctgaata gcatatatat gatgcacogg ataggtcatt atgatttttt accatttcga 1860
cttacataat gaaaaccaat tcattttaaa taccagatta ttattttgta agttgtggaa 1920
aaagctaatt gtagttttca ttatgaagtt ttcccaataa accagggtatt ctaaaaaaaa 1980
aaaaaaaa

```

&lt;210&gt; 174

&lt;211&gt; 238

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 174

```

Gly Ala Ala Ser Pro Arg Pro Leu Arg Phe Cys Gly Gly Ala Arg Ala
 1           5           10           15
Arg Arg Pro Leu Ser Ala Val Ala Arg Pro Ala Arg Ser Ser Asp Pro
          20          25          30
Leu Arg Ser Ala Pro Leu Gly Pro Ala Pro Pro Val Asn Met Ile Arg
          35          40          45
Cys Gly Leu Ala Cys Glu Arg Cys Arg Trp Ile Leu Pro Leu Leu Leu
 50          55          60
Leu Ser Ala Ile Ala Phe Asp Ile Ile Ala Leu Ala Gly Arg Gly Trp
 65          70          75          80
Leu Gln Ser Ser Asp His Gly Gln Thr Ser Ser Leu Trp Trp Lys Cys
          85          90          95
Ser Gln Glu Gly Gly Gly Ser Gly Ser Tyr Glu Glu Gly Cys Gln Ser
          100         105         110
Leu Met Glu Tyr Ala Trp Gly Arg Ala Ala Ala Met Leu Phe Cys
          115         120         125
Gly Phe Ile Ile Leu Val Ile Cys Phe Ile Leu Ser Phe Phe Ala Leu
          130         135         140
Cys Gly Pro Gln Met Leu Val Phe Leu Arg Val Ile Gly Gly Leu Leu
145          150         155         160
Ala Leu Ala Ala Val Phe Gln Ile Ile Ser Leu Val Ile Tyr Pro Val
          165         170         175
Lys Tyr Thr Gln Thr Phe Thr Leu His Ala Asn Pro Ala Val Thr Tyr
          180         185         190
Ile Tyr Asn Trp Ala Tyr Gly Phe Gly Trp Ala Ala Thr Ile Ile Leu
          195         200         205
Ile Gly Cys Ala Phe Phe Phe Cys Cys Leu Pro Asn Tyr Glu Asp Asp
210          215         220
Leu Leu Gly Asn Ala Lys Pro Arg Tyr Phe Tyr Thr Ser Ala
225          230         235

```

&lt;210&gt; 175

&lt;211&gt; 4181

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 3347, 3502, 3506, 3520, 3538, 3549, 3646, 3940, 3968, 3974,

4036, 4056, 4062, 4080, 4088, 4115  
 <23> n = A,T,C or G

<400> 175

```

ggtggatgcg tttgggtgt agctaggctt ttttttttt tttcttttta aaacacatct 60
agacaaggaa aaaaacaagc tccgatctga tttttcactc ctcgttcttg tctgtggttc 120
ttactgtggt tgtgtatttt aaaggcgaga agacgagggg aacaaaacca gctggatcca 180
tccatcaccc tgggtgggtt taatttttct ttttttctcg ttattttttt ttaaacacac 240
actcttcaca atgaacaaac tgtatatcgg aaacctcagc gagaacgcog cccctcggga 300
cctagaaagt atcttcaagg acgccaagat cccggtgtcg ggacccttcc tggtagaagc 360
tggctacgct tctgtggact gcccggaaga gagctgggac ctcaaggcca tcgaggcgct 420
ttcaggttaa atagaactgc acgggaaacc catagaagtt gagcactcgg tcccaaaaag 480
gcaaaagatt cggaaacttc agatacgaag tatcccgctc catttacagt gggagggtgt 540
ggatagttaa ctagtccagt atggagtggt ggagagctgt gagcaagtga acactgactc 600
ggaaactgca gttgtaaagt taacctatc cagtaaggac caagctagac aagcactaga 660
caaatggaat ggattttcagt tagagaattt caccttgaag ctgagctata tccctgatga 720
aatgctgcgc cagcaaaaacc ccttgcagca gccccgaggt cgccgggggc ttgggcagag 780
gggtcctcca agcagcgggt ctccaggatc cgtatccaag cagaaaaact gtgattttgc 840
tctgcgctgt cgtgttccca cccaatttgt tggagccatc ataggaaaag aggtgtgccac 900
cattcggaac atccaaaaac agaccocgat taaaatcgat gtccaccgta aagaaaatgc 960
ggggctgctc gagaagtcca ttactatcct ctctactcct gaaggcacct ctgcggcttg 1020
taagtctatt ctggagatta tgcataaaga agctcaagat ataaaattca cagaagagat 1080
ccctctgaag attttagctc ataataaact tgttggacgt ctctattgta agaaaggaa 1140
aaatcttaaa aaaattgagc aagacacaga cactaaaatc acgatattcc cattgcagga 1200
attgacgtgt tataatccag aacgcactat tacagttaaa ggcaatgttg agacattgac 1260
caagctgtag gaggagatca tgaagaaaat cagggagctc atgaaaaatg atattgcttc 1320
tatgaatctt caagcacatt taattctctg attaaatctg aacgccttgg gtctgttccc 1380
accacttcaa gggattccac ctcccactc agggccctct tcacgatga ctctccctca 1440
cccgcaattt gagcaatcag aaacggagac tgttcatcag tttatccca ctctatcagt 1500
cggctgcact atcgggaacg agggccagac catcaagcag ctctctcgct ttgctggagc 1560
ttcaattaaag attgctccag cggaagcacc agatgctaaa gtgaggatgg tgattatcac 1620
tggaaccaaca gaggctcagt tcaaggctca gggagaattt tatggaaaaa ttaaaagaca 1680
aaaacttggt agtctaaag aagaggtgaa acttgaagct catctcagag tgccactcct 1740
tgctgctggc agagttattg gaaaaggagg caaaacgggt aatgaacttc agaattttgc 1800
aagtgcagaa gttgtttgct ctctgaccca gacacctgat gagaatgacc aagtggttgt 1860
caaaataaact ggtcactttc atgcttgcca ggttgccagc agaaaaatc aggaattctc 1920
gactcaggta aagcagcacc aacaacagaa ggctctgcaa agtggaaccac ctcagctcag 1980
acggaagttaa aggtcaggga aacagccacc cacagaggca gatgccaaac caaagacaga 2040
ttgcttaacc aacagatggg cgtcgacccc ctatccagaa tcacatgcac aagtttttcc 2100
ctagocagct gttctgaggc accaggcaac ttttgaaact ctgtctctgt gagaatgtat 2160
actttatgct ctctgaaagt tatgacaccc agcttttaaa caaacaaaac aacaaaacaa 2220
aaaaagggtg gggaggaggg gaaagagaag agctctgca tttcccttgg tgtagtctca 2280
cagtaataca gatattctaa ttcttcttaa tattccccc taatgccaga aattggotta 2340
atgatgcttt cactaaatc atcaaataga ttgctcctaa atccaattgt taaaattgga 2400
tcagaaatct tatcacagga acttaaatgt taagccatta gcataacaaa actgttctca 2460
gtttttattt tacctaacc taacatgagt aacctaaagg aagtgtgaa tgggtgtggc 2520
aggggtgatt aacgtcattt ttactcaac taacctcaggt attcagtaat acgaatgaaa 2580
gcaaaattgt tctctttttt tgaaaaattt atatacttta taatgataga agtccaaccc 2640
ttttttaaaa aataaattta aaattttaaa gcaatcagct aacaggcaaa ttaagatttt 2700
tactcttgcc tgggtgacagt aaaagctgaa aattaatttc aggggttttt gagggttttg 2760
acaacagtat tagttaaact aaatgttcaa aaatacggag cagtgctcag tatctggaga 2820
gcagcatctac catttatctc ttcatattata gttgggaaag tttacgagc tactaacaaa 2880
tgggtcgcag gagatttttg aacggctcgt ttaaatgctt tcaggagact ctgatttttt 2940
gtttagctac atgattgaat gcataataaa ttctgactat caatacctaa 3000
agaaagtgca tcagtgaaag gatgcaagac tttaactgta ctggcaaaaa gcaagcttta 3060
gottgtctta taggatgctt agtttgccac tactactcag accaatggga cagtcataga 3120
tgggtgaca gtgtttaaac gcaacaaaag gctacatttc catggggcca cactgtcat 3180
gagctcactc aagctatttt gaagattttt aagcactgat aaattaaaaa aaaaaaaaaa 3240

```

```

aaattagact ccaccttaag tagtaaagta taacaggatt ttctgtatact gtgcaatcag 3300
ttctttgaaa aaaagtcaaa aagatagaga atacaagaaa agttttnggg atataatttg 3360
aatgactgtg aaaacatatg acctttgata acgaactcat ttgtctcactc cttgacagca 3420
aagcccagta cgtacaattg tggtgggtgt ggggtgtctc caaggccaag ctgctctctg 3480
aattgatttt ttgagttttg gnttgnaaga tgatcacagn catgttaaac tgatcttnaa 3540
ggacatatnt tataacocctt taataacocctt atccocctgcc tcattcttat ttcgagatga 3600
atttcgatgc agactagatg tctttctgaa gatcaattag acattntgaa aatgatttaa 3660
agtgttttcc ttaatgttct ctgaaaacaa gtttcttttg tagttttaac caaaaaagtg 3720
ccctttttgt cactggtttc tcttagcatt catgattttt ttttcacaca atgaattaaa 3780
attgctaaaa tcatggactg gctttctggt tggatttcag gtaagatgtg ttaaggcca 3840
gagcttttct cagtatttga ttttttccc caatatttga ttttttaaaa atatacacat 3900
aggagctgca tttaaaacct gctggtttaa attctgtcan atttcacttc tagcctttta 3960
gtatggcnaa tcanaattta cttttactta agcatttgta atttggagta tctggtagta 4020
gctaagaat aattcnataa ttgagttttg tactcnccaa anatgggtca ttcctcatgn 4080
ataatgtncc cccaatgcag cttcattttc caganacctt gacgcaggat aaattttttc 4140
atcatttagg tcccaaaaaa aaaaaaaaaa aaaaaaaaaa a 4181

```

&lt;210&gt; 176

&lt;211&gt; 579

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 176

```

Met Asn Lys Leu Tyr Ile Gly Asn Leu Ser Glu Asn Ala Ala Pro Ser
1 5 10 15
Asp Leu Glu Ser Ile Phe Lys Asp Ala Lys Ile Pro Val Ser Gly Pro
20 25 30
Phe Leu Val Lys Thr Gly Tyr Ala Phe Val Asp Cys Pro Asp Glu Ser
35 40 45
Trp Ala Leu Lys Ala Ile Glu Ala Leu Ser Gly Lys Ile Glu Leu His
50 55 60
Gly Lys Pro Ile Glu Val Glu His Ser Val Pro Lys Arg Gln Arg Ile
65 70 75 80
Arg Lys Leu Gln Ile Arg Asn Ile Pro Pro His Leu Gln Trp Glu Val
85 90 95
Leu Asp Ser Leu Leu Val Gln Tyr Gly Val Val Glu Ser Cys Glu Gln
100 105 110
Val Asn Thr Asp Ser Glu Thr Ala Val Val Asn Val Thr Tyr Ser Ser
115 120 125
Lys Asp Gln Ala Arg Gln Ala Leu Asp Lys Leu Asn Gly Phe Gln Leu
130 135 140
Glu Asn Phe Thr Leu Lys Val Ala Tyr Ile Pro Asp Glu Met Ala Ala
145 150 155 160
Gln Gln Asn Pro Leu Gln Gln Pro Arg Gly Arg Arg Gly Leu Gly Gln
165 170 175
Arg Gly Ser Ser Arg Gln Gly Ser Pro Gly Ser Val Ser Lys Gln Lys
180 185 190
Pro Cys Asp Leu Pro Leu Arg Leu Leu Val Pro Thr Gln Phe Val Gly
195 200 205
Ala Ile Ile Gly Lys Glu Gly Ala Thr Ile Arg Asn Ile Thr Lys Gln
210 215 220
Thr Gln Ser Lys Ile Asp Val His Arg Lys Glu Asn Ala Gly Ala Ala
225 230 235 240
Glu Lys Ser Ile Thr Ile Leu Ser Thr Pro Glu Gly Thr Ser Ala Ala
245 250 255
Cys Lys Ser Ile Leu Glu Ile Met His Lys Glu Ala Gln Asp Ile Lys
260 265 270
Phe Thr Glu Glu Ile Pro Leu Lys Ile Leu Ala His Asn Asn Phe Val

```





&lt;213&gt; Homo sapiens

&lt;400&gt; 178

```

agcgcctttca aggggtgtacg caaagcactc attgataccc ttttgatgg ctatgaaaca 60
gccgcgtatg ggacaggggt ctttggccag aatgagtagc tacgctatca ggaggccctg 120
agttagctgg ccaactgcggg taaagcacga attgggagct ctacagcgaca tcaccagtca 180
gcagccaaag acctaaactca gtccctcgag gtctcccca caaccatcca ggtgacatca 240
ctcccctcca gtcagaagag taaacgtgcc aagcacttcc ttgaattgaa gagctttaag 300
gataactata acacattgga gagtactctg tgacggagct gaaggactct tgcgtagat 360
taagccagtc agttgcaatg tgcaagacag gctgcttgcc gggccgcctt cggaaacatct 420
ggcccagcag gccacagact tatccatcca agttccggtt gtatccagag ttcttagagc 480
ttgtgtctaa agggtaattc cccaaccctt ccttatgagc atttttagaa cattggctaa 540
gactattttc cccagtagcg g

```

&lt;210&gt; 179

&lt;211&gt; 521

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 179

```

cccaacgcgt ttgcaaatat tcccctggta gctactcttc ttacccccga atattggtaa 60
gatcgagcaa tggccttcagg acatgggttc tcttctctcg tgatcatcca agtgctcact 120
gcacgaagac tggctttgtct cagtgtttca acctaccag ggctgtctct tggccacac 180
ctcgctccct gttagtcccg tatgacagcc cccatcaaat gacottggcc aagtcaacgt 240
ttctctgtgg tcaagggttg ttggctgatt ggtggaagat agggtgacc aaaggaggcc 300
acgtgagcag tcagcaaccag ttctgcacca gcagcgccct cgtcctagtg ggtgttctcg 360
ttctctctgg ccttgggtgg gctagggcct gattcgggaa gatgcctttg cagggagggg 420
aggataagtg ggatctacca attgattctg gcaaaacatt ttctaagatt tttttgcttt 480
atgtgggaaa cagatctaaa tctcatttta tgotgtattt t

```

&lt;210&gt; 180

&lt;211&gt; 417

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 180

```

gggtggaatc gccgaagatg gcggaggtgc aggtctcgtt gcttgatggt cgaggccatc 60
tctggggcgg cctggcgccc atcgtggcta aacaggtagt gctggggcgg aaggtgtgtg 120
tgtagcgtct tgaaggcatc aacatttctg gcaatttcta cagaacaag ttgaagtacc 180
tggctttctc cgcgaagcgg atgaacacca accctcccg agggccctac cacttccggg 240
ccccagcccg catcttctgg cggaccgtgc gaggtatgct gccccacaaa accaagcgag 300
gccaggccgc totgacgcgt ctcaagggtt ttgacggcat cccacgcgcc tacgacaaga 360
aaaagcggat ggtggttctt gctgccctca aggtcgtcgg tctgaagcct acaagaa 417

```

&lt;210&gt; 181

&lt;211&gt; 283

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 35

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 181

```

gatttctctt aaataggatg taaaacttct ttcanattac tcttctctcag tcttgcctgc 60
caagaactca agtgaactcg tgataaata acctttccca ggtatattgg caggtatgtg 120
tgtaattctca gaatacacag gtgacataga tatgatatga caactggtaa tgggtgattc 180

```

```

atttacattg tttacacttc tatgaccagg ccttaaggga aggtcagttt tttaaaaaac 240
caagtagtgt ctctctacct atctccagat acatgtcaaa aaa 283

```

```

<210> 182
<211> 401
<212> DNA
<213> Homo sapiens

```

```

<400> 182
atattcttgc tgccttatgca gctgacattg ttgccctccc taaagcaacc aagtagcctt 60
tatttcccac agtgaaagaa aacgctggcc tatcagttac attacaaaag gcagatttca 120
agagagattga gtaagtagtt ggatggcctt cataaaaaa agaattcaag aagaggattc 180
atgctttaag aaacatttgt tatacattcc tcacaaatta tacctgggat aaaaactatg 240
tagcaggcag tgtgttttcc ttccatgtct ctctgcaact cctgcagttg gtccctctag 300
gotgcaagtc tgtcctatct gaattcccag cagaagcact aagaagctcc accctatcac 360
ctagcagata aaactatggg gaaaacttaa atctgtgcat a 401

```

```

<210> 183
<211> 366
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 325
<223> n = A,T,C or G

```

```

<400> 183
accgtgtcca agttttttag acccttggtta gccagaccga ggtgtcctgg tcaccgtttc 60
accatcatgc tttgatgttc cccgtgtctt ctctctctcg ctctcaagag caaagggttaa 120
tttaaggaca aagatgaagt cactgtaaac taatctgtca ttgtttttac ctctcctttc 180
tttttcagtg cagaaattaa aagtaagtat aaagcacogt gattgggagt gtttttgcgt 240
gtgtcgggat cactggtaaa tgttggctga gaacaatccc tccccttgca ctgtgaaaa 300
cactttgagc gctttaagag attancctga gaaataatta aatatctttt ctcttcaaaa 360
aaaaaa 366

```

```

<210> 184
<211> 370
<212> DNA
<213> Homo sapiens

```

```

<400> 184
tcttacttca aaagaaaaat aaacataaaa aataagttgc tggttcctaa caggaaaaat 60
tttaataatt gtactgagag aaactgctta cgtacacatt gcagatcaaa tatttggagt 120
taaaatgtta gtctacatag atgggtgatt gtaactttat tgccattaaa agatttcaaa 180
ttgcatctcat gctctctgtg acacataaat aaaaatgggc aaataatgaa gatctctcct 240
tcagctctgt ctgtttaatt ctcttctcta atgtcggctc cctaattgta 300
cacagtttag tgatatctag gagtataaag ttgtcgccca tcaataaaaa tcacaaagtt 360
ggtttaaaaa 370

```

```

<210> 185
<211> 107
<212> DNA
<213> Homo sapiens

```

```

<400> 185
ctcatattat tttccttttg agaaattgga aactctttct gttgctatta tattaataaa 60
gttggtgttt attttctggt agtcaccttc ccattttaa aaaaaaa 107

```

&lt;210&gt; 186

&lt;211&gt; 309

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 186

gaaaggatgg	ctctgtgtgc	cacagagctg	ggacttcatg	ttcttctaga	gagggccaca	60
agagggccac	aggggtggcc	gggagtgtgc	agctgatgcc	tgctgagagg	caggaattgt	120
gccagtga	gacagtcatg	agggagtgtc	tcttcttggg	gaggaagaa	ggtagagcct	180
ttctgtctga	atgaaaggcc	aaggctacag	tacagggccc	cgccccagcc	aggggtgttaa	240
tgcccacgta	gtggaggcct	ctggcagatc	ctgcattcca	aggtcactgg	actgtacgtt	300
tttatggtt						309

&lt;210&gt; 187

&lt;211&gt; 477

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 187

ttcagtccta	gcaagaagcg	agaattctga	gatcctccag	aaagtgcagc	agcaccacc	60
tcacaacctcg	ggccagtgtc	ttcaggcttt	actggggacc	tgcgagctgg	cctaattgtg	120
tgccctctcaa	gccaggccat	ccctggggcg	cacagacgag	ctccgagcca	ggtcaggctt	180
cgaggggccac	aagctcagcc	tcaggccccc	gcactgattg	tgccagaggg	gccactaccc	240
aaggtctagc	tagggccaag	acctagttaa	ccagacagtg	agaagccccc	ggaaggcaga	300
aaagtgtgga	gcattggcaga	caggggaagg	aaacattttc	agggaaaaga	catgtatcac	360
atgtcttcag	aagcaagtca	ggtttcatgt	aacogagtg	ctctcttgct	gtccaaaagt	420
agcccagggc	tgtagcacag	gcttcacagt	gattttgtgt	tcagccgtga	gtcacac	477

&lt;210&gt; 188

&lt;211&gt; 220

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 188

ttaatatggt	agatattaat	attcctctta	gatgaccagt	gattccaatt	gtcccaagtt	60
ttaaataagt	accctgtgag	tatgagataa	attagtgaca	atcagaacaa	gtttcagtat	120
cagatgttca	agaggaagtt	gctattgcat	tgattttaat	atttgtacat	aaacactgat	180
ttttttgagc	attattttgt	atttgttgta	ctttaataac			220

&lt;210&gt; 189

&lt;211&gt; 417

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc feature

&lt;222&gt; 76, 77

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 189

accatcttga	cagaggatac	atgctcccaa	aacgtttgtt	accacactta	aaaatcactg	60
ccatcatata	gcacatnntt	caaaattata	gccattcatg	atttactttt	tcagatgac	120
tatcattatt	ctagtctctt	gaatttggta	ggggaaaaaa	aacaaaaaca	aaaacttacg	180
atgcactttt	ctccagcaca	tcagatttca	aattgaaaaa	taaagacatg	ctatggtaat	240
gcaattgcta	gtactacaca	ctttgtacaa	caaaaaacag	aggcaagaaa	caacggaag	300
agaaaaagc	tcctttgttg	gcccttaaac	tgagtcaaga	tctgaaatgt	agagatgatc	360
tctgacgata	cctgtatgtt	cttattgtgt	aaataaaatt	gctggtatga	aatgaca	417

<210> 190  
 <211> 497  
 <212> DNA  
 <213> Homo sapiens

<400> 190  
 gcactgcggc gctctccgt cccgcggtgg ttgtgtctgc tgcgcgtgct gctgggctg 60  
 aacgcaggag ctgtcattga ctggcccaca gaggaggcca aggaagtatg ggattattgt 120  
 acggtccgca aggatgccta catgttcttg ttgctctatt atgccaccaa ctctgtcaag 180  
 aacttctcag aactgccctt ggatcatgtg cttcaggcggt gtccaggcgg ttctagcact 240  
 ggatttgaaa aactttgagg aattggggccc cttgacagt atctcaaaacc acgaaaaacc 300  
 acctggctcc aggcctgccag tctcttattt gtggataatc ccgtgggcac tgggttcagt 360  
 tatgtgaatg gtatgtgtgc ctatgccaa gacctggcta tgggtggctt agacatgatg 420  
 gttctctgta agacctctt cagtggccac aaagaattcc agacagttcc attctacatt 480  
 ttctcagagt cctatgg 497

<210> 191  
 <211> 175  
 <212> DNA  
 <213> Homo sapiens

<400> 191  
 atgttgaaat ttttgcttat taactttgtt tattgtcttc tcctctgatt agaattattg 60  
 ctacttgagt acaaggattt gagcctgtta cattcactgc tgaatttttag gctcctggaa 120  
 gataccagc attcaatga gaccacacaa taaatatatg tcaataaaaa aaaaa 175

<210> 192  
 <211> 526  
 <212> DNA  
 <213> Homo sapiens

<400> 192  
 agtaaacatt attatTTTT ttatatttgc aaaggaaaca tatctaattc ttctataga 60  
 aagaacagta ttgctgtaat tctttttctt ttcttctca ttctctgc cccttaaaag 120  
 attgaagaaa gagaaacttg tcaactcata tccaogttat ctagcaaaagt acataaagaat 180  
 ctatcactaa gtaatgtatc cttcagaatg tgttggttta ccagtgcac ccctattca 240  
 tcacaaaatt aaagcaagaa gtccatagta atttatttgc taatagtggg ttttataatg 300  
 tcagagtctc tgagggtcaa ttttatcttt tcaactcaaa gctctatgat cttaataaat 360  
 ttacttaagt tatttttggt tattttcttc aaataaatat tgggtgtcaa gactatatct 420  
 aattcctctg atcactttga gaaacaaact tttattaaat gtaaggcact tttctatgaa 480  
 ttttaatat aaaaaataat attgttctga ttattactga aaaaa 526

<210> 193  
 <211> 553  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 290, 300, 411, 441  
 <223> n = A,T,C or G

<400> 193  
 tccattgtgg tggaattcgc tctctggtaa aggcgtgcag gtgttgccg cggtctctga 60  
 gctgggatga gccgtgctcc cgttggaagc aaggagccc agccggagcc atggccagta 120  
 cagtggtagc agttggactg accattgctg ctgcaggatt tgcaggcgt tacgttttgc 180  
 aagccatgaa gcatatggag cctcaagtaa aacaagtttt tcaaaagccta ccaaaatctg 240

```

ccttcagtgg  tggctattat  agaggtgggt  ttgaacccaa  aatgacaaan  cggggaagcan  300
cattaatact  aggtgtaagc  cctactgcc  ataaagggaa  aataagagat  gctcatcgac  360
gaattatgct  tttaaatcat  cctgacaaag  gaggatctcc  ttatatagca  nccaaaaatca  420
atgaagctaa  agatttacta  naagggtcaag  ctaaaaaatg  aagtaaatgt  atgatgaatt  480
ttaagtctgt  attagtttat  gtatatgagt  actaatgttt  tataataaaa  tgcctcagag  540
ctacaatttt  aaa

```

&lt;210&gt; 194

&lt;211&gt; 320

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 194

```

cccttcccaa  tccatcagta  aagaccccat  ctgccttgtc  catgcccgtt  cccaacaggg  60
atgtcacttg  atatgagagt  ctcaaatctc  aatgccttat  aagcattcct  tccgtgtgac  120
attaagactc  tgataattgt  ctccctcca  taggaatttc  tcccaggaaa  gaaatatatc  180
cccatctccg  tttcatatca  gaactaccgt  ccccgatatt  cccttcagag  agattaaaga  240
ccagaaaaaa  gtgagcctct  tcatctgcac  ctgtaatagt  ttcagttcct  attttcttcc  300
attgacccat  atttatacct

```

&lt;210&gt; 195

&lt;211&gt; 320

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 203, 218

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 195

```

aagcatgacc  tggggaaatg  gtcagacctt  gtattgtgtt  tttggccttg  aaagtagcaa  60
gtgaccagaa  tctgccatgg  caacaggctt  taaaaaagac  ccttaaaaag  acactgtctc  120
aactgtgggt  ttgacaccag  ccagctctct  gtacatttgc  tagctttag  ttttctaaga  180
ctgagtaaac  ttcttatttt  tanaaagggg  aggcgtgntt  gtaactttcc  ttgacttaaa  240
ttgggtaaaa  gtcttttcca  caaaccacca  tctattttgt  gaactttggt  agtcatcttt  300
tatttggtaa  attatgaact

```

&lt;210&gt; 196

&lt;211&gt; 357

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 36

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 196

```

atataaaata  atacgaaact  ttaaaaagca  ttggantgtc  agtatgttga  atcagtagtt  60
tcaactttaac  tgtaaaacaat  ttcttaggac  accatttggg  ctagtttctg  tgtaagtgtg  120
aatactacaa  aaacttattt  atactgttct  tatgtcattt  gttatattca  tagattttata  180
tgatgatatg  acatctggct  aaaaagaagt  tattgcaaaa  ctaaccacta  tgactttttt  240
tataaatact  gtatggacaa  aaaaatggcat  tttttatatt  aaattgttta  gctctggcaa  300
aaaaaaaaaa  ttttaagagc  tgggtactaat  aaaggattat  tatgactgtt  aaaaaaa  357

```

&lt;210&gt; 197

&lt;211&gt; 565

<212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 27  
 <223> n = A,T,C or G

<400> 197  
 tcagctgagt accatcagga tatttanccc ttaagtgtct gttttgggag tagaaaaacta 60  
 aagcaacaat acttccctct gacagctttg attggaatgg gggtattaga tcattcaacct 120  
 tggctctaca ctttttagga tgcttgggtga acataacacc acttataatg aacatccctcg 180  
 gtctctatat ttgggctat gtgggttagga attgttactt gttactgcag cagcagccct 240  
 agaaagtaag ccagggcct cagatctaag ttagtccaaa agctaaatga tttaaagtca 300  
 agttgtaatg ctaggcataa gcaactctata atacattaaa ttataggcgg agcaattagg 360  
 gaatgtttct gaacatttaa acttgtattt atgtcactaa aattctaaca caaacttaaa 420  
 aaatgtgtct catacatatg ctgtactagg ctctcatatg cattttctaaa tttgtgtatg 480  
 atttgaatat atgaagaagt ttatacaaga gtgtatttta aaattattaa aaataaatgt 540  
 atataatttg tacctattgt aaaaa 565

<210> 198  
 <211> 484  
 <212> DNA  
 <213> Homo sapiens

<400> 198  
 tatgtaagta ttggtgtctg ctttaaaaaa ggagaccocag acttccacctg tccttttttaa 60  
 acatttgaga acagtgttac tctgagcagt tggggcaacct tcaccttatc cgacagctga 120  
 ctgttggatg tgtccattgt cgccagtttg gctgttgccc ggacaggaca ggacctccat 180  
 tgggcgcagc agcaggtggc aggggtgtgg cttgaggtgg gtggcagcgt ctggtccctcc 240  
 tctctgtgtc tttctgagag ggtctctaaa gcagagtgtg gttggcctgg ggggaaggcag 300  
 agcacgtatt tctccctct agtacctctg catttgtgag tgttccctct ggcctttctga 360  
 agggcagcag actcttgagt atactgcaga ggacatgctt tatcagtagg tcctgagggc 420  
 tccaggggct caactgacca agtaacacag aagttggggg atgtggccta tttgggtcgg 480  
 aaac

<210> 199  
 <211> 429  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 77, 88, 134, 151, 189, 227, 274, 319  
 <223> n = A,T,C or G

<400> 199  
 gcttatgttt tttgttttaa cttttgttt ttaacattta gaatattaca ttttgtatta 60  
 tacagtacct ttctcanaca tttgtanaa ttcatttcgg cagctcaacta ggaattttgtc 120  
 gaacattaaa aagngtgata gcgatattag ngccaatcaa atggaaaaaa ggtagtctta 180  
 ataaacaana cacaacgttt ttatacaaca tacttttaaaa tattaanaaa actccttaat 240  
 attgtttcct attaagtatt attccttggg caanattttc tgatgotttt gattttctct 300  
 caatttagca tttgctttnng gttttttct ctatttagca ttctgttaag gcacaaaaac 360  
 tatgtactgt atgggaaatg ttgtaaatat taccttttcc acattttaaa cagacaactt 420  
 tgaatccaa 429

<210> 200  
 <211> 279

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 200

```

gcttttttga ggaattacag ggaagctcct ggaattgtac atggatatct ttatccctag 60
ggggaaatca aggagctggg caccctaata tctttatgga agtggttaaa actattttaa 120
ttttattaca agtattacta gagtatgtgt tctactctaa gatttcaaaa gtgcatttaa 180
aatcatacat gtccccgcct gcaaatata tctttatttg gtggagaaaa aaatagtata 240
ttctacataa aaaattaaag atattaacta agaaaaaaa                279

```

&lt;210&gt; 201

&lt;211&gt; 569

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 201

```

taggtcagta tttttagaaa ctcttaatag ctcatactct tgataccaaa agcagccctg 60
attgtttaaa cacacacctg cacaagaagc agtcatgggt gcatttacct ttctctgggtg 120
cacaaaaaaa aattctcaaa aagcaaggac ttacgctttt tgcaaaagcct ttgagaagtt 180
actggatcat aggaagctta taacaagaat ggaagattct taataaactc actttctttg 240
gtatccagta acagtagatg ttcaaaatat gtagctgatt aataccagca ttgtgaacgc 300
tgtacaacct ttgggttatt actaagcaag ttactactag ctctctgaaa gtacgttcat 360
aattaatggt atttatacac tgccctccat gacttttact ttgccctaag ctaatctcca 420
aaatctgaaa tgctactcca atatcagaaa aaaaggggga ggtggaatta tatttctcgt 480
gattttaaga gtacagagaa tcatgcacat ctctgattag tcatatatg tctagtgtgt 540
aataaaagtc aaagatgaac tctcaaaaa                569

```

&lt;210&gt; 202

&lt;211&gt; 501

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 202

```

attaatagcc ttaataattg ttggcaagga tccctttgct tcttttgcca tgcaagctcc 60
tagcatctgg cagtggggcc aagaaaaata ggtttatgca tgtatgatgg ttttctttct 120
gagcaacatg attgagaacc agtgtatgtc aacaggtgca tttgagataa ctttaaatga 180
tgtacctgtg tggctaaagc ttggaatctgg tcaaccttca tccatgcaac aacttgttca 240
aattcttgac aatgaaatga agctcaatgt gcatatggat tcaatccac accatcgatc 300
atagaccacc ctatcagcac tgaaaaactct ttgcatataa gggatcattg caagagcagc 360
gtgactgaca ttatgaagcc ctgtactgaa gacagcaagc tgttagtaga gaccagatgc 420
ttctctggca ggctcgttgt acctcttgga aaacctcaat gcaagatagt gtttcagtgc 480
tggcatattt tggaaattctg c                501

```

&lt;210&gt; 203

&lt;211&gt; 261

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 36, 96

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 203

```

gacaagctcc tggctcttgag atgtcttctc gtttaangaga tgggcctttt ggaggtaaa 60
gataaaatca atgagttctg tcatgattca ctattntata acttgcatga cctttactgt 120
gttagctcct tgaatgttct tgaaatttta gactttcttt gtaaacaaat gatagtgtct 180
tatcattgta taaaagctgt tatgtgcaac agtgtggaga ttcttctgtc gatttaataa 240

```

## 100

aataacttaaa cactgaaaaa a

261

&lt;210&gt; 204

&lt;211&gt; 421

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 204

```

agcatctttt ctacaacgtt aaaattgcag aagtagctta tcattaaaaa acaacaacaa 60
caacaataac aataaatcct aagtgtaaat cagttattct accccctacc aaggatatca 120
gcctgttttt tcoccttttt ctocctgggaa taattgtggg ctctctccca aatttctaca 180
gcctcttttc tcttctcatg cttagagcttc cctgttttgc cgcattcggtg tgcaggagctg 240
gcttgtgtgc ttggactcgg ctccaggtgg aagcatgctt tcccttggtta ctgttggaga 300
aactcaaac ttcaagccct aggtgtagcc attttgtcaa gtatcaact gtatttttgt 360
actggcatta acaaaaaaag aagataaaat attgtaccat taaactttta taaaacttta 420
a

```

421

&lt;210&gt; 205

&lt;211&gt; 460

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 205

```

tactctcaca atgaaggacc tgggaatgaaa aatctgtgtc taaacaagtc ctcttttagat 60
tttagtgcaa atccagagcc agcgtcggtt gcctcgagta attctttcat gggtaaccttt 120
ggaaagctc tcaggagacc tcacctagat gcctattcaa gctttggaca gccatcagat 180
tgtcagccaa gagcctttta tttgaaagct cattcttccc cagacttggg ctctgggtca 240
gaggaagatg ggaagaaaag gacagatttt caggaagaaa atcacatttg tacccttttaa 300
cagactttag aaaactacag gactocaaat ttccagttct atgacttggg cacatagact 360
gaatgagacc aaaggaaaag cttaacatac tacotcaagg tgaactttta tttaaaagag 420
agagaatctt atgtttttta aatggagtta tgaattttta

```

460

&lt;210&gt; 206

&lt;211&gt; 481

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 206

```

tgtggtggaa ttccgggacgc cccagagccc tgacttttct ctgcgtgggc cgtctctctc 60
tgccgaagca gtgacctctg acccctggtg acctctgctt tgagtgcctt ttgaacgctg 120
gtcccgcggtg acttgggttt ctcaagctct gtctgtccaa agacgctccg gtccaggtcc 180
cgctgcctcc gggtggatcc ttgaacccca gacgcccctc tgtgtgctgt tgtccggagg 240
cggtccttcc atctgcctgc ccaccggag ctctttccgc cggcgcaggg tcccagccc 300
acctcccgc ctcagtcctg cgggtgtggt ctgggacagt cctgcacaca caatgcaagt 360
cetgtgcctcc gcgcccgccc gccgtaccgc ccgccaactc tgttatttat 420
ggtgtgaccc cctggagggtg cccctcgccc accggggcta tttattgttt aatttatttt 480
t

```

481

&lt;210&gt; 207

&lt;211&gt; 605

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 207

```

accctttttg gattoagggc tcttcacaat taaaatgagt gtaatgaaac aagggtaaaa 60
tatagaagca tccctttgta tactgttttg ctacttacag tgtacttggc attgctttat 120
ctcactggat tctcacggta ggattttctg atacttaac taagctccaa agttgtctac 180
ttttttgata ctagggtgct cctttttgtt tacagagcag ggtcacttga tttgtcagct 240

```



```

ggtggcagaa ttggcaccat tacccaggto tgactgacca ccagtcagag gcactttatt 300
tgtatcatga aatgatttga aatcatttga aagcagcgaa gtctgataat gaatgccagc 360
tttccttgtg ctttgataac aaagactcca aatattctgg agaacctgga taaaagtttg 420
aagggttaga ttgggatttg aagacaaaat tgtaggaaat cttacatttt tgcaataaca 480
aacattaatg aaagcaaaac attataaaag taattttaat tcaccacata cttatcaatt 540
tcttgatgct tccaaatgac atctaccaga tatggttttg tggacatctt tttctgttta 600
cataa 605

```

&lt;210&gt; 208

&lt;211&gt; 655

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 208

```

ggcggttgct tggattcccg tcgttaactta aagggaacct ttccacaatgt ccggagccct 60
tgatgtcctg caaatgaagg agggaggatgt ccttaagttc cttgcagcag gaacccactt 120
aggtggcacc aatcttgact tccagatgga acagtacatc tataaaaggga aaagtgtatg 180
catctatctc ataaatctca agaggacctg ggagaagctt ctgctggcag ctogtgcgaat 240
tgttgccatt gaaaacctcg ctgatgtcag tgttatatcc tccaggaata ctggccagag 300
ggctgtgctg aagtttgctg ctgccactgg agccaactcca attgctggcc gcttcaactc 360
tggaaacctc actaaaccaga tccaggcagc cttccgggag ccacggcttc ttgtgggttac 420
tgacccacag gctgaccacc agcctctcac ggaggcatct tatgttaacc tacctaccat 480
tgcgctgtgt aacacagatt ctctctgcg ctatgtggac attgccatcc catgcaacca 540
caagggaagt caactcagtg gtttgatgtg gtggatgctg gtcgggaag ttctgcgcac 600
gcgtggcacc atttcccgta aacacccatg ggaggtcatg cctgatctgt acttc 655

```

&lt;210&gt; 209

&lt;211&gt; 621

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 209

```

catttagaac atggttatca tccaagacta ctctaccctg caacattgaa ctcccaagag 60
caaatccaca ttctcttga gttctgcagc ttctgtgttaa atagggcagc tgtctgtctat 120
gccgtagaat cacatgatct gaggaccatt catggaagct gctaaaatgc ctagtctggg 180
gagctctcca taaagttttg catggagcaa acaaacagga ttaaaactagg tttggttctc 240
tcagccctct aaaaacatag ggcctagcct gcaggcttcc ttgggctttc tctgtgtgtg 300
tagttttgta aacactatag catctgttaa gatccagtgt ccatggaaac cttccacatc 360
gccgtgactc tggactatat cagtttttgg aaagcagggt tctctgctct gctaaacagc 420
ccacgtggac cagctcgaat gtctttcctt tacacctatg tttttaaatc gtcacaaactc 480
aagaaaaaat ctaaacagat ttctgttgca tatgtgtttg tgaactgtga tttgtattta 540
gtaggctctc atattgcatt taacttgttt ttgtaactcc tgattcttcc ttttoggata 600
ctattgatga ataagaaat t 621

```

&lt;210&gt; 210

&lt;211&gt; 533

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 20, 21, 61

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 210

```

cgccctgggg agccggcggn ngagtcgggg acgtggagac ccgggggtccc ggcagccggg 60
ngcccgggg gccacgggtg gggatgcacc gcccggggtt gggagctggc gccatcgcca 120
agaagaaact tgcagaggcc aagtataaag agcgagggac ggtcttggct gaggaccagc 180

```

```

tagccagat gtcaaagcag ttggacatgt tcaagaccaa cctggaggaa tttgccagca 240
aacacaagca ggagatccgg aagaatcctg agttccgtgt gcagttccag gacatgtgtg 300
caaccattgg cgtggatccg ctggcctctg gaaaaggatt ttggtctgag atgctgggag 360
tgggggacct ctattacgaa ctagggtgtcc aaattatcga agtgtgcctg gcgctgaagc 420
atcggaatgg aggtctgata acctttggagg aactacatca acaggtgttg aagggaaggg 480
gcaagttcgc ccaggatgtc agtcaagatg acctgatcag agcoatcaag aaa 533

```

```

<210> 211
<211> 451
<212> DNA
<213> Homo sapiens

```

```

<400> 211
ttagcttgag ccgagaacga ggcgagaaag ctggagaccg aggagaccgc ctagagcgga 60
gtgaacgggg aggggaccgt ggggaccgcg ttgatcgtgc gcggacacct gctaccaagc 120
ggagcttcag caaggaaagt gaggagcgga gttagagaaag gccctccag cctgaggggc 180
tgcccaaggc agctagcctc acggaggatc gggaccgtgg gcgggatgcc gtgaagcgag 240
aagctgcctc acccccagtg agcccccctga aggcggctct ctctgaggag gagttagaga 300
agaaatccaa ggctatcatt gaggaatato tccatctcaa tgacatgaaa gaggcagttc 360
agtgcgtgca ggagctggcc tcaccctcct tgcctctcat ctttgaacgg catggtgtgc 420
agtctacgct ggagcgagct gccattgtgc g

```

```

<210> 212
<211> 471
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 54
<223> n = A,T,C or G

```

```

<400> 212
gtgattatcc ttgatcaggg agaagatcat ttgattttgt ttgcattcc ttanaatgga 60
gggcaacatt ccacagctgc cctggctgtg atgagtgtcc ttgcaggggc cggagtagga 120
gcactggggg gggggcgga ttggggttac tcgatgtaag ggattccttg ttgttgtgtt 180
gagatccagt gcagtttgtga tttctgtgga tccagccttg gttccaggaa ttttgtgtga 240
ttggcttaaa tccagtttcc aatctctgac agctgggctg gaactggaac tcagtagctg 300
aaactgtctg acccggtcac gttcttgat cctcagaact ctttgcctct gtcggggtgg 360
gggtgggaac tcacgtgggg agcgggtggc gagaaaatgt aaggattctg gaatacatat 420
tccatgggac tttccttccc tctcctgctt cctcttttcc tgcctccata c 471

```

```

<210> 213
<211> 511
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 27, 63, 337, 442
<223> n = A,T,C or G

```

```

<400> 213
ctaattagaa acttgcgtga ctttttnttt tcttttaggg gtcaaggacc ctctttatag 60
ctnccatttg cctacaataa attattgcag cagtttgcaa tactaaaata ttttttatag 120
aactttatatt ttttcctttg ataaagggat gctgcatagt agagttgggt taattaaact 180
atctcagcgc tttccctgtc ttccctctgc ctccatagc ctcatgtgoc ttcaggagag 240
ctcttttaat cttaaagttc tacatttcat gctcttagtc aaattctggt acctttttaa 300

```

```

taactcttcc cactgcata ttcacatcttg aattgngngt tctaaattct gaaactgtag 360
ttgagatata gctatttaaat atttctggga gatgtgoatc cctcttcttt gtgggtgcc 420
aaggttggtt tgcgtaactg anactccttg atatgottca gagaatttag gcaaacactg 480
gccatggcgg tgggagtact gggagtaaaa t 511

```

&lt;210&gt; 214

&lt;211&gt; 521

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 214

```

agcattgcc aataatccct aattttccac taaaaatata atgaaatgat gttaaagctt 60
ttgaaaagtt taggttaaac ctactgttgt tagattaatg tatttggttg ttccctttat 120
ctggaaatgtg gcattagcct ttttatttta accctcttta attcttattc aattccatga 180
cttaagggtt gagagctaaa cactggggatt ttgggataac agactgacag ttttgcatata 240
ttataatcgg cattgtacat agaaaggata tggctacctt ttgttaaatc tgcactttct 300
aaatatcaaa aaaggggaat gaagtataaa tcaatttttg tataatctgt ttgaaacatg 360
agttttattt gcttaataatt agggcctttgc cctttttctg taagtctctg gggatcctgt 420
gtagaagctg ttctcatata acaccaaaca gttaagtcca ttctctgcta ctagtacaa 480
attcggtttc atattctact taacaattta aataaactga a 521

```

&lt;210&gt; 215

&lt;211&gt; 381

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 17, 20, 60, 61, 365

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 215

```

gagcggagag cggaacngtn agagccctga gcagccccc cgcgcgcgc gcgcctagttn 60
ncatcacacc cggggaggag ccgcagctgc cgcagccgcg ccaggtcacc atcacccgcaa 120
ccatgagcag cgagggccgag acccagcagc cgcgcccgcc ccccccgcgc gcccccgcgc 180
tcagcgcgcg cgacaccaag cccggcacta cgggcagcgg cgcagggagc ggtggcccg 240
cggccctcac atcggcggcg cctgcggcg gggacaagaa ggtcatcgca acgaaggttt 300
tgggaacagt aaaatggttc aatgtaagga acggatattg ttcatcaac aggaattgaca 360
ccaangaaga tgtatttgta c 381

```

&lt;210&gt; 216

&lt;211&gt; 425

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 216

```

ttactaacta ggtcattcaa ggaagtcaag ttaacttaaa catgtcacct aaatgcactt 60
gatggtgttg aaatgtccac cttcttaaat ttttaagatg aacttagttc taaagaagat 120
aacaggccaa tcttgaaggt actccctgtt tgctgcagaa tgtcagatat tttgtagtt 180
gcataaagagt cctatttgcc ccagttaatt caacttttgt ctgacctgtt tgtggactgt 240
ctggctctgt tagaactctg tccaaaaagt gcattgata taacttgtaa agcttcccac 300
aattgacaat atatatgcac gtgttttaaac caaatccaga aagcttaaac aatagagctg 360
cataatagta tttattaaag aatcacacact gtaaacatga gaataactta aggattctag 420
tttag 425

```

&lt;210&gt; 217

&lt;211&gt; 181

&lt;212&gt; DNA

<213> Homo sapiens

<400> 217

```
gagaaaccaa atgataggtt gtagagcctg atgactccaa acaaagccat caccgcgatt 60
cttcctcctt cttctggtgc tacagctcca agggcccttc accttcatgt ctgaaatgga 120
actttggcgt tttcagtgtg agaataatgt gaagggtttca tttgtttcta gaaaaaaaaa 180
a 181
```

<210> 218

<211> 405

<212> DNA

<213> Homo sapiens

<400> 218

```
caggccttcc agttcactga caaacatggg gaagtgtgcc cagctggctg gaaacctggc 60
agtataacca tcaagcctga tgtccaaaag agcaaagaat atttctccaa gcagaagtga 120
gcgctgggct gtttttagtgc caggctgcgg tgggcagcca tgagaacaaa acctcttctg 180
tatttttttt ttccattagt aaaacacaag acttcagatt cagcogaatt gtggtgtctt 240
acaaggcagg cctttcctac aggggggtgga gagaccagcc ttctctcctt tggtaggaat 300
ggcctgagtt ggctgtgtgt gcaggtctact ggtttgtatg atgtattagt agagcaaccc 360
attaatcttt tgtagtttgt attaaacttg aactgagaaa aaaaa 405
```

<210> 219

<211> 216

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 207, 210

<223> n = A,T,C or G

<400> 219

```
actccaagag ttagggcagc agagtggagc gatttagaaa gaacatttta aaacaatcag 60
ttaatttacc atgtaaaatt gctgtaaatt ataattgtga cagattttct gtccaatat 120
tcaattgtaa actcttgggt aagactgtta cgtttctatt gcttttgtat gggatattgc 180
aaaaataaaa aggaagaagc cctcttnaan aaaaa 216
```

<210> 220

<211> 380

<212> DNA

<213> Homo sapiens

<400> 220

```
cttacaattt gccccatgtt gtaggggaca cagaaccttt tgagaaaact tagatttttt 60
cttgtaacaa gtctttgcct ttttctctct tcattttttt ccagtacatt aaatttgta 120
atttcattct tgagggaacac tgattagatg ggtgtgtgtt gtgtttgatg ggagaaaaca 180
gcacccaag gactcagaag atgattttta cagttcagaa cagatgtgtg caattattgt 240
gcattgaata atgttgagt gcagtcacaaa gtcattgatt ttactcttag tcttcattac 300
tgcattgaaa aggaaaacct gtctgagaaa atgcctgaca gtttaattta aaactatggt 360
gtaagtcttt gacaaaaaaa 380
```

<210> 221

<211> 398

<212> DNA

<213> Homo sapiens

<400> 221

```

ggttagtaag ctgtcgactt tgtaaaaaag ttaaaaaatga aaaaaaaagg aaaaatgaat 60
tgtatatatta atgaatgaac atgtacaatt tgcactcagg aggaggttcc tttttgttg 120
gtgagtcctgc aagtgaattt cactgatgtt gatattcatt gtgtgtagtt ttatttcggt 180
ccacgcccgc tttcctttta ttttgagctt aatgccagct gcgtgtctag ttttgagtcg 240
agtaaaatag aatcagcaaa tcaactctat ttttcatcct tttccggtat tttttggggt 300
gtttctgtgg gagcagtgta caccaactct tccgttatat tgcccttttg ctggaaaatg 360
ttgtatgttg aataaaattt totataaaaa ttaaaaaa 398

```

```

<210> 222
<211> 301
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 49, 64
<223> n = A,T,C or G

```

```

<400> 222
ttcgataaatt gatctcatgg gctttccctg gaggaagggt ttttttgnr gtttttttt 60
taanaaccttg aaactttgtaa actgagatgt ctgtagcttt ttggcccatc tgtagtgtat 120
gtgaagattt caaaacctga gagcactttt tctttgttta gaattatgag aaaggcacta 180
gatgacttta ggattttgcat ttttcccttt attgcctcat ttctgtgac gcottgttg 240
ggagggaat ctgtttattt tttctactaa ataaaaagct aagattctat atcgcaaaaa 300
a 301

```

```

<210> 223
<211> 200
<212> DNA
<213> Homo sapiens

```

```

<400> 223
gtaagtgcctt aggaagaaac tttgcaaaac tttaatgagg atacctgtt cattttttaa 60
attccttcac actgtaattt aatgtgtttt atattctttt gtagtataac aacataaact 120
agatttctac aggagacagt ggtttttatt ggattgtctt ctgtaatagg tttcaataaa 180
gctggatgaa cttaaaaaaa 200

```

```

<210> 224
<211> 385
<212> DNA
<213> Homo sapiens

```

```

<400> 224
gaaaggtttg atccggactc aaagaaagca aaggagtgtg agccgccatc tgctggagca 60
gctgttaactc caagacctgg acaagagatt cgtcagcgaa ctgcagctca aagaacatt 120
tctccaacac cagcaagccc taaccagggc cctcctccac aagttccagt atctcctgga 180
ccacaaagg acagttctgc cctgggtgga cccacagaaa ggactgttac tccagcccta 240
tcataaatg tgttaccaag acatcttgga tccctgcta cttcagtgcc tggaaatgggt 300
aaacagagca cttaagtgtt ttacagttt atattgttt ctctggttac caataaaacg 360
ggccattttc agtggtgtaa aaaaa 385

```

```

<210> 225
<211> 560
<212> PRT
<213> Homo sapiens

```

```

<400> 225
Met Glu Cys Leu Tyr Tyr Phe Leu Gly Phe Leu Leu Leu Ala Ala Arg

```

1	5	10	15
Leu Pro Leu Asp	Ala Ala Lys Arg	Phe His Asp Val	Leu Gly Asn Glu
20		25	30
Arg Pro Ser Ala	Tyr Met Arg Glu His	Asn Gln Leu Asn	Gly Trp Ser
35		40	45
Ser Asp Glu Asn	Asp Trp Asn Glu	Lys Leu Tyr Pro	Val Trp Lys Arg
50		55	60
Gly Asp Met Arg	Trp Lys Asn Ser	Trp Lys Gly Arg	Val Gln Ala
65		70	75
Val Leu Thr Ser	Asp Ser Pro Ala	Leu Val Gly Ser	Asn Ile Thr Phe
	85	90	95
Ala Val Asn Leu	Ile Phe Pro Arg	Cys Gln Lys Glu	Asp Ala Asn Gly
100		105	110
Asn Ile Val Tyr	Glu Lys Asn Cys	Arg Asn Glu Ala	Gly Leu Ser Ala
115		120	125
Asp Pro Tyr Val	Tyr Asn Trp Thr	Ala Trp Ser Glu	Asp Ser Asp Gly
130		135	140
Glu Asn Gly Thr	Gly Gln Ser His	His Asn Val Phe	Pro Asp Gly Lys
145		150	155
Pro Phe Pro His	His Pro Gly Trp	Arg Arg Trp Asn	Phe Ile Tyr Val
	165	170	175
Phe His Thr Leu	Gly Gln Tyr Phe	Gln Lys Leu Gly	Arg Cys Ser Val
180		185	190
Arg Val Ser Val	Asn Thr Ala Asn	Val Thr Leu Gly	Pro Gln Leu Met
195		200	205
Glu Val Thr Val	Tyr Arg Arg His	Gly Arg Ala Tyr	Val Pro Ile Ala
210		215	220
Gln Val Lys Asp	Val Tyr Val Val	Thr Asp Gln Ile	Pro Val Phe Val
225		230	235
Thr Met Phe Gln	Lys Asn Asp Arg	Asn Ser Ser Asp	Glu Thr Phe Leu
	245	250	255
Lys Asp Leu Pro	Ile Met Phe Asp	Val Leu Ile His	Asp Pro Ser His
260		265	270
Phe Leu Asn Tyr	Ser Thr Ile Asn	Tyr Lys Trp Ser	Phe Gly Asp Asn
275		280	285
Thr Gly Leu Phe	Val Ser Thr Asn	His Thr Val Asn	His Thr Tyr Val
290		295	300
Leu Asn Gly Thr	Phe Ser Leu Asn	Leu Thr Val Lys	Ala Ala Ala Pro
305		310	315
Gly Pro Cys Pro	Pro Pro Pro Pro	Pro Pro Pro Pro	Pro Ser Lys Pro
	325	330	335
Pro Ser Leu Gly	Pro Ala Gly Asp	Asn Pro Leu Glu	Leu Ser Arg Ile
340		345	350
Pro Asp Glu Asn	Cys Gln Ile Asn	Arg Tyr Gly His	Phe Gln Ala Thr
355		360	365
Ile Thr Ile Val	Glu Gly Ile Leu	Glu Val Asn Ile	Ile Gln Met Thr
370		375	380
Asp Val Leu Met	Pro Val Pro Trp	Pro Glu Ser Ser	Leu Ile Asp Phe
385		390	395
Val Val Thr Cys	Gln Gly Ser Ile	Pro Thr Glu Val	Cys Thr Ile Ile
	405	410	415
Ser Asp Pro Thr	Cys Glu Ile Thr	Gln Asn Thr Val	Cys Ser Pro Val
420		425	430
Asp Val Asp Glu	Met Cys Leu Leu	Thr Val Arg Arg	Thr Phe Asn Gly
435		440	445
Ser Gly Thr Tyr	Cys Val Asn Leu	Thr Leu Gly Asp	Asp Thr Ser Leu
450		455	460
Ala Leu Thr Ser	Thr Leu Ile Ser	Val Pro Asp Arg	Asp Pro Ala Ser

```

465                      470                      475                      480
Pro Leu Arg Met Ala Asn Ser Ala Leu Ile Ser Val Gly Cys Leu Ala
                      485                      490                      495
Ile Phe Val Thr Val Ile Ser Leu Leu Val Tyr Lys Lys His Lys Glu
                      500                      505                      510
Tyr Asn Pro Ile Glu Asn Ser Pro Gly Asn Val Val Arg Ser Lys Gly
                      515                      520                      525
Leu Ser Val Phe Leu Asn Arg Ala Lys Ala Val Phe Phe Pro Gly Asn
                      530                      535                      540
Gln Glu Lys Asp Pro Leu Leu Lys Asn Gln Glu Phe Lys Gly Val Ser
545                      550                      555                      560

```

```

<210> 226
<211> 9
<212> PRT
<213> Homo sapiens

```

```

<400> 226
Ile Leu Ile Pro Ala Thr Trp Lys Ala
1                      5

```

```

<210> 227
<211> 9
<212> PRT
<213> Homo sapiens

```

```

<400> 227
Phe Leu Leu Asn Asp Asn Leu Thr Ala
1                      5

```

```

<210> 228
<211> 9
<212> PRT
<213> Homo sapiens

```

```

<400> 228
Leu Leu Gly Asn Cys Leu Pro Thr Val
1                      5

```

```

<210> 229
<211> 10
<212> PRT
<213> Homo sapiens

```

```

<400> 229
Lys Leu Leu Gly Asn Cys Leu Pro Thr Val
1                      5                      10

```

```

<210> 230
<211> 10
<212> PRT
<213> Homo sapiens

```

&lt;400&gt; 230

Arg Leu Thr Gly Gly Leu Lys Phe Phe Val  
 1 5 10

&lt;210&gt; 231

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 231

Ser Leu Gln Ala Leu Lys Val Thr Val  
 1 5

&lt;210&gt; 232

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 232

Ala Gly Ala Asp Val Ile Lys Asn Asp Gly Ile Tyr Ser Arg Tyr Phe  
 1 5 10 15  
 Phe Ser Phe Ala  
 20

&lt;210&gt; 233

&lt;211&gt; 21

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 233

Phe Phe Ser Phe Ala Ala Asn Gly Arg Tyr Ser Leu Lys Val His Val  
 1 5 10 15  
 Asn His Ser Pro Ser  
 20

&lt;210&gt; 234

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 234

Phe Leu Val Thr Trp Gln Ala Ser Gly Pro Pro Glu Ile Ile Leu Phe  
 1 5 10 15  
 Asp Pro Asp Gly  
 20

&lt;210&gt; 235

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 235

Leu Gln Ser Ala Val Ser Asn Ile Ala Gln Ala Pro Leu Phe Ile Pro



109

1 5 10 15  
 Pro Asn Ser Asp  
 20

<210> 236  
 <211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 236  
 Ile Gln Asp Asp Phe Asn Asn Ala Ile Leu Val Asn Thr Ser Lys Arg  
 1 5 10 15  
 Asn Pro Gln Gln  
 20

<210> 237  
 <211> 21  
 <212> PRT  
 <213> Homo sapiens

<400> 237  
 Arg Asn Ser Leu Gln Ser Ala Val Ser Asn Ile Ala Gln Ala Pro Leu  
 1 5 10 15  
 Phe Ile Pro Pro Asn  
 20

<210> 238  
 <211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 238  
 Thr His Glu Ser His Arg Ile Tyr Val Ala Ile Arg Ala Met Asp Arg  
 1 5 10 15  
 Asn Ser Leu Gln  
 20

<210> 239  
 <211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 239  
 Arg Asn Pro Gln Gln Ala Gly Ile Arg Glu Ile Phe Thr Phe Ser Pro  
 1 5 10 15  
 Gln Ile Ser Thr  
 20

<210> 240  
 <211> 21  
 <212> PRT  
 <213> Homo sapiens

&lt;400&gt; 240

Gly	Gln	Ala	Thr	Ser	Tyr	Glu	Ile	Arg	Met	Ser	Lys	Ser	Leu	Gln	Asn
1				5					10					15	
Ile	Gln	Asp	Asp	Phe											
				20											

&lt;210&gt; 241

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 241

Glu	Arg	Lys	Trp	Gly	Phe	Ser	Arg	Val	Ser	Ser	Gly	Gly	Ser	Phe	Ser
1				5					10					15	
Val	Leu	Gly	Val												
				20											

&lt;210&gt; 242

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 242

Gly	Ser	His	Ala	Met	Tyr	Val	Pro	Gly	Tyr	Thr	Ala	Asn	Gly	Asn	Ile
1				5					10					15	
Gln	Met	Asn	Ala												
				20											

&lt;210&gt; 243

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 243

Val	Asn	His	Ser	Pro	Ser	Ile	Ser	Thr	Pro	Ala	His	Ser	Ile	Pro	Gly
1				5					10					15	
Ser	His	Ala	Met												
				20											

&lt;210&gt; 244

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 244

Ala	Val	Pro	Pro	Ala	Thr	Val	Glu	Ala	Phe	Val	Glu	Arg	Asp	Ser	Leu
1				5					10					15	
His	Phe	Pro	His												
				20											

&lt;210&gt; 245

&lt;211&gt; 20

&lt;212&gt; PRT

<213> Homo sapiens

<400> 245

Lys Pro Gly His Trp Thr Tyr Thr Leu Asn Asn Thr His His Ser Leu  
 1 5 10 15  
 Gln Ala Leu Lys  
 20

<210> 246

<211> 20

<212> PRT

<213> Homo sapiens

<400> 246

Asn Leu Thr Phe Arg Thr Ala Ser Leu Trp Ile Pro Gly Thr Ala Lys  
 1 5 10 15  
 Pro Gly His Trp  
 20

<210> 247

<211> 20

<212> PRT

<213> Homo sapiens

<400> 247

Leu His Phe Pro His Pro Val Met Ile Tyr Ala Asn Val Lys Gln Gly  
 1 5 10 15  
 Phe Tyr Pro Ile  
 20

<210> 248

<211> 20

<212> PRT

<213> Homo sapiens

<400> 248

Pro Glu Thr Gly Asp Pro Val Thr Leu Arg Leu Leu Asp Asp Gly Ala  
 1 5 10 15  
 Gly Ala Asp Val  
 20

<210> 249

<211> 20

<212> PRT

<213> Homo sapiens

<400> 249

Gly Phe Tyr Pro Ile Leu Asn Ala Thr Val Thr Ala Thr Val Glu Pro  
 1 5 10 15  
 Glu Thr Gly Asp  
 20

<210> 250

<211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 250  
 Phe Asp Pro Asp Gly Arg Lys Tyr Tyr Thr Asn Asn Phe Ile Thr Asn  
 1 5 10 15  
 Leu Thr Phe Arg  
 20

<210> 251  
 <211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 251  
 Leu Gln Ala Leu Lys Val Thr Val Thr Ser Arg Ala Ser Asn Ser Ala  
 1 5 10 15  
 Val Pro Pro Ala  
 20

<210> 252  
 <211> 153  
 <212> PRT  
 <213> Homo sapiens

<400> 252  
 Met Ala Ser Val Arg Val Ala Ala Tyr Phe Glu Asn Phe Leu Ala Ala  
 1 5 10 15  
 Trp Arg Pro Val Lys Ala Ser Asp Gly Asp Tyr Tyr Thr Leu Ala Val  
 20 25 30  
 Pro Met Gly Asp Val Pro Met Asp Gly Ile Ser Val Ala Asp Ile Gly  
 35 40 45  
 Ala Ala Val Ser Ser Ile Phe Asn Ser Pro Glu Glu Phe Leu Gly Lys  
 50 55 60  
 Ala Val Gly Leu Ser Ala Glu Ala Leu Thr Ile Gln Gln Tyr Ala Asp  
 65 70 75 80  
 Val Leu Ser Lys Ala Leu Gly Lys Glu Val Arg Asp Ala Lys Ile Thr  
 85 90 95  
 Pro Glu Ala Phe Glu Lys Leu Gly Phe Pro Ala Ala Lys Glu Ile Ala  
 100 105 110  
 Asn Met Cys Arg Phe Tyr Glu Met Lys Pro Asp Arg Asp Val Asn Leu  
 115 120 125  
 Thr His Gln Leu Asn Pro Lys Val Lys Ser Phe Ser Gln Phe Ile Ser  
 130 135 140  
 Glu Asn Gln Gly Ala Phe Lys Gly Met  
 145 150

<210> 253  
 <211> 462  
 <212> DNA  
 <213> Homo sapiens

<400> 253  
 atggccagtg tccgcgtggc ggccactctt gaaaaacttc tcgcggcggtg gcggcccgtg 60

## 113

aaagcctctg	atggagatta	ctacaccttg	gctgtaccga	tgggagatgt	accaatggat	120
ggatctctctg	tgtctgatat	tggagcagcc	gtctctagca	tttttaattc	tccagaggaa	180
tttttaggca	aggccgtggg	gctcagtgca	gaagcactaa	caatacacga	atatctgcat	240
gtttttcca	aggctttggg	gaaagaagtc	cgagatgcga	agattaccoc	ggaagctttc	300
gagaagctgc	gattccctgc	agcaaaaggaa	atagccaata	tgtgtcgttt	ctatgaaatg	360
aagcgcagcc	gagatgtcaa	tctcaccocac	caactaaatc	ccaaagtcaa	aagcttcacg	420
cagtttatct	cagagaacca	gggagccttc	aagggcattg	ag		462

&lt;210&gt; 254

&lt;211&gt; 8031

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 254

tggcgaaatg	gacgcgccct	gtagcggcgc	attaagcgcg	gcggggtgtg	tggttacgcg	60
cagcgtgacc	gctacacttg	ccagcgccct	agcgcgccct	cctttcgctt	tcttcccttc	120
ctttctcgcc	acgttcgcgc	gctttccccc	tcaagctcta	aatcgggggc	tcoccttagg	180
gttccgattt	agtgctttac	ggcacctcga	cccaaaaaaa	ctgtattagg	gtgatgggtc	240
acgtagtggg	ccatcgccct	gatagacggt	ttttcgccct	ttgacgttgg	agtcacagctt	300
ctttaatagt	ggactcttgt	tccaaactgg	aacaacactc	aacctatctc	cgggtctatc	360
ttttgattta	taagggattt	tgocgatttc	ggcctattgg	ttaaaaaaat	agctgattta	420
acaaaaattt	aacgcgaatt	ttacaaaaat	attaacgttt	acaatttcag	gtggcaactt	480
tccgggaaat	gtgcgcggaa	ccctattttg	tttatttttc	taaatcacatt	caaatatgta	540
tcgctcatcg	aattaattct	tagaaaaact	catcgagcat	caaatgaaac	tgcaatttat	600
tcatatcagg	attatcaata	ccatattttt	gaaaaagccg	ttctgtaat	gaaggagaaa	660
actcaccgag	gcagttccat	aggatggcaa	gactcctgta	tcgggtctgc	attccgactc	720
gtccaaacatc	aatacaacct	attaattttc	cctcgtcaaa	aataagggtta	tcaagtfgaa	780
aatcaccatg	agtgacagct	gaatccggtg	agaatggcaa	aagtttatgc	attttcttcc	840
agacttgttc	aacagccag	ccattacgct	cgtcatcaaa	atcactogca	tcaacaaac	900
cgttattcat	tctgtatgtc	gcctgagcga	gacgaaatca	gcgatcgctg	ttaaaaaggac	960
aattacaaac	aggaatcgaa	tgcaaccgcg	gcaggaaacac	tgccagcgca	tcaacatat	1020
tttcacctga	atcaggatat	tcttctaata	cctggaaatg	tgttttcccg	gggatcgacg	1080
tggtagataa	ccatgcatac	tcaggagtac	ggataaaatg	cttgatggtc	ggaagaggca	1140
taaatccgtt	cagccagttt	agtctgacca	tctcatctgt	aacatcatgt	gcaacgctac	1200
ctttgccatg	tttcagaaac	aactctggcg	catcgggctt	cccatacaat	cgatagattg	1260
tcgcacgaat	ttgcccgaca	ttatcgcgag	cccatattta	cccatataaa	tcagcatcca	1320
tgttggaatt	taatcgcgcg	ctagagcaag	ccttttcccg	ttgaatatgg	ctcataaac	1380
cccttgtatt	actgtttatg	taagcagaca	gttttattgt	tcatgaccaa	aatcccttaa	1440
cgtgagtttt	cgttcacagt	agcgtcagac	cccgtagaaa	agatcaaaag	gtcttcttga	1500
ctttgcctttt	ttctcgcgct	aatctcgtcg	ttgcacaaaca	aaaaaccacc	atccatcagc	1560
gtggtttgtt	tgccggatca	agagctacca	actctttttc	cgaaggttaa	tggtctcagc	1620
agagcgcaga	taccaaatac	tgtccttcta	gtgtagccgt	agttaggcca	ccaactcaag	1680
aactctgtag	caaccgcctac	atacctcgct	ctgctaatac	tgttaccagt	ggctgctgcc	1740
agtgggcaga	agtcgtgtct	taccgggttg	gactcaagac	gatagttacc	ggataaggcg	1800
cagcgcgtgg	gctgaacgcg	gggttcgtgc	acagagccca	gcttggagcg	aacgacctac	1860
accgaactga	gataacctaca	gcgtgagcta	tgagaaagcg	ccacgctctc	cgaagaggta	1920
aagcgcgcaga	ggtaacctgc	aagcgcgca	gtcggaaacag	gagagcgcac	gagggagcgt	1980
ccagggggaa	acgccttgta	tctttatagt	cctgtccggt	ttcgccacct	ctgacttgag	2040
cgtcgatttt	tgtgatgctc	gtcagggggg	cgagccttat	ggaaaaaccg	cagcaacgcg	2100
gcctttttac	gggttcctg	cttttctgct	ccctttgctc	acatgtttct	tctctgggtta	2160
tcctcctgatt	ctgtggataa	ccgtattacc	gaccttgagt	gagctgatac	cgctcgccgc	2220
agccgaacga	cogagccggt	cgagtcatgt	agcgaagagc	cggaaagagc	ctgtatgcgc	2280
tattttctcc	ttacgcatct	gtgcggtatt	tcacaccgca	tatatgtgtg	actctcagta	2340
caatctgctc	tgtatgcgca	tagttaagcc	agtatacact	ccgctatcgc	taactgactg	2400
ggctcatggct	gcgcccgcac	accgcacca	accgcgtgac	gcgcctgac	gggcttgtct	2460
gtccccgcga	tccgcttaca	gacaagctgt	gaccgtctcc	gggagctgca	tgtgtcagag	2520
gttttccacg	tcatcacgga	aacgcgcgag	gcagctcgcg	taaaagctcat	acagctgtgtc	2580
gtgaagcgat	tcaagatgtt	ctcgcgtgtc	atccgcgtcc	agctcgttga	gtttctccag	2640

aagcgcttaat	gtctggcttc	tgataaagcg	ggccatgtta	agggcggttt	tttctgttt	2700
ggtoactgat	gcoctcoggt	aagggggatt	tctgttcoat	ggggtaatga	taccgatgaa	2760
acgagagagg	atgctcacga	tacgggttac	tgatgatgaa	catgccccgt	tactggaaag	2820
ttgtgagggt	aaacaactcg	cggatggat	gcggcgggac	cagagaaaaa	tcactcaggg	2880
tcaatgccag	cgcttcgtta	atacagatgt	aggtgttcca	cagggttagcc	agcagcatcc	2940
tgcgatcgag	atccggaaca	taatggtgca	ggcgctgac	ttcccgcttt	ccagacttta	3000
cgaaacacgc	aaaccogaaga	coattcatgt	tgttgcctag	ctgcogaagc	ttttgcagca	3060
gcagtgcgtt	cacgttcgct	cgcgtatcgg	tgattcattc	tgctaaccag	taaggcaacc	3120
cgcgcagcct	agccgggttc	tcaacgacag	gagcacgcatc	atgcgcacc	atgcgcgcgc	3180
catgcgcggc	ataatggcoct	gottctcgcc	gaacgctttg	gtggcgggac	cagtgacgaa	3240
ggcttgagcg	agggcggtgca	agattccgaa	taccgcaagc	gacaggccga	tcactcgtcg	3300
gctccagcga	aagcggctcct	cgccgaaaat	gaccagagac	gctgcggcca	cctgtcctac	3360
gagttgcag	ataaagaaga	cagtcataag	tcggcgcgacg	atagtcatgc	cccgccgcca	3420
coggaaggag	ctgactgggt	tgaaggctct	caagggcctc	ggtcgagatc	ccggtgccta	3480
atgagttagc	taacttacat	taattgcgtt	gcgctcactg	cccgctttcc	agtcgggaaa	3540
cctgtcgtgc	cagctgcatt	aatgaatcgg	ccaacgcgcg	gggagagcgg	gitttgcgtat	3600
tgggcgccag	gggtggtttt	cttttcacca	gtgagcaggg	caacagctga	ttgcctctca	3660
ccgctggggc	ctgagagagt	tgacgcaagc	ggtccacgct	ggtttgcccc	agcaggcgcaa	3720
aatcctgttt	tagggtggtt	aaocggcgga	tataacatga	gctgtcttcg	gtatcgtcgt	3780
atcccactac	cgagatatcc	gcaccaacgc	gcagcccgga	ctcggttaatg	gcgcgcatgt	3840
cgcccgacgc	cactcgatcg	ttggcaacca	gcctgcagtg	gggaacgatg	ccctaatcca	3900
gcaattcgag	ggtttggtag	aaaccggaca	tggcactcca	ctgcgcttcc	cgttcgccta	3960
tcggctgaat	ttgattgoga	tgagatatatt	tatgccagcc	agccagacgc	agacgcgcgc	4020
agacagaact	taattggggcc	gctaacaagc	cgatttgcgt	gtgaccaaat	gcgaccagat	4080
gctccaagcc	cagtcgcgta	cogtcttcat	gggagaaaaat	aatactgttg	attgggtgtct	4140
ggtcagagac	atcaagaagt	aaocggcgaa	cattagtcca	ggcagcttcc	acagcaatgg	4200
catctcgttc	atccagcgga	tagttaatga	tcagcccaat	gacgcgttgc	gcggaagaat	4260
tgctcaccgc	cgctttacag	gcttcgacgc	cogcttcgtt	taccactcgc	accaccaagc	4320
tggaaccagc	tgattcggcg	cgagatttaa	tcgcgcgcag	aatttgcgac	gcgcgcttga	4380
gggcacagat	ggaggtggca	acgccaatca	gcacagactg	tttgcgcgcg	agttgtttgtg	4440
ccacgcggtt	gggaatgtaa	ttcagctcgc	ccatcgcgcg	ttccactttt	tcocgcgttt	4500
tcgcagaaac	gtggctggcc	tggttcacca	cgcgggaaac	ggctctgataa	gagacaccgc	4560
catactctgc	gacatcgat	aaogttactg	gtttcacatt	ccacaccctg	aattgactct	4620
cttcgcggcg	ctatcatgac	ataccgcgaa	aggttttgcg	ccattcgatg	gtgtccggga	4680
tctcgacgct	ctcccttagt	cgactcctgc	attaggaagc	agcccgatag	taggttgagg	4740
cogtttgagca	cgcgcgcgcg	aaggatgtgt	gcagtcaagc	agatggcgcc	caacagctcc	4800
cgcgcacagc	ggcctgcacc	cataccacag	ccgaacaacg	cgctcatgac	ccogaagtgg	4860
cgagcccgat	cttccccatc	gggtgatgtg	gcgatatagg	cgccagcaac	cgccactgtg	4920
gcgcgcggtga	tgccggccac	gatgcgttcg	gcgtagagga	tcgagatctc	tgaccocgga	4980
aattaatacag	actcaactata	ggggaattgc	gagcggataa	caattccctc	ctagaataata	5040
ttttgtttaa	ctttaagaag	gagatataca	tatgcagcat	caccaccact	accacggagt	5100
acagcttcaa	gacaatgggt	ataatggatt	gctcatttga	attaactctc	aggttaactga	5160
gaatcagaac	ctcatctcaa	acattaagga	aatgataaact	gaagcttcat	tttaactatt	5220
taatgctcac	aagagaagag	tatttttcaag	aaataataag	attttaatac	ctgcaccaatg	5280
gaaagcttaac	aataacagca	aaataaaaaa	gaagagcgaa	atgtcatagt	gtactcatgt	5340
gactgactgc	tatggggcac	atgggagatga	tcacataccc	ctacaataga	gaggggtgtg	5400
aaaagaggga	aaatacattc	atttccaccc	taatttctca	ctgaattgata	acttaacagc	5460
tggctacgga	tcacgaggcc	gagtgtttgt	ccatgaatgg	gcccacotcc	gttggggtgt	5520
gttcgagtga	tataacaagt	acaaaccttt	ctacataaat	gggcaaaatc	aatttaaagt	5580
gacaaggtgt	tcacttgaca	tcacaggcat	ttttgtgtgt	gaaaagggtc	cttgcccccga	5640
agaaaaactgc	attattagta	agcttttttaa	agaaggatgc	acotttatct	acaatagcac	5700
ccaaaaatgca	actgcataca	taatgttcat	gcgaagtgtta	tcctctgtgg	ttgaattttg	5760
taatgcaagt	accaccaacc	aagaagcacc	aaacctacag	aaccagatgt	gcagcctcag	5820
aagtgcactgg	gatgtataca	cagactctgc	tgactttcac	ccagcgtttc	ccatgaacgg	5880
gactgagctt	ccacctctc	ccacattctc	gcttgtagag	gctggtgaca	aagtgtgtctg	5940
tttagtgctg	gatgtgtcca	gcaagatggc	agaggctgac	agactccttc	aactacacaa	6000
agccgcagaa	ttttatttga	tcagatttgt	tgaattcatg	acottgtgtg	gacttgccag	6060
tttcgacagc	aaaggagaga	tcagagccca	gctacaccaa	ataacagca	atgatgatgc	6120

```

aaagttgctg gtttcatatc tgcccaccac tgtatcagct aaaaacagaca tcagcatttg 6180
ttcaggggctt aagaaaggat ttgaggtggg tgaaaaactg aatggaaaag cttatggctc 6240
tgtgatgata ttagtaccac gccggagatga taagctctctt ggcaatttgt taccoactgt 6300
gctcagcagat gggtcaacaa ttcaactccat tgccctgggt tcactcgcag ccccaaatct 6360
ggaggaaatta tcaagtcctta caggaggttt aaagttcttt gttccagata tatcaaacctc 6420
caatagcatg attgatgctt tcagtagaat ttctctgga actggagaca ttttccagca 6480
acataattcag ctgaaagta caggtgaaaa tgtcaaacct caccatcaat tgaaaaacac 6540
agtgaactg gataatactg tgggcaacga cactatgttt ctagttaagt ggcaggccag 6600
tggctctctt gagattatat tatttgatcc tgaatggaaga aaatactaca caaataattt 6660
tatcaccaat ctaacttttc ggacagctag tctttggatt ccaggaaacag ctaagccctg 6720
gcaactggact tacacccctga acaataccca tcattctctg caagccctga aagtgcactg 6780
gaacctctgc gccccaact cagctgtgcc cccagccact gtggaagcct ttgtggaag 6840
agacagcctc caatttctct atcctgtgat gatttatgcc aatgtgaac agggatttta 6900
tccattctt aatgccactg tcaactgccac agttgagcca gagactggag actctgttac 6960
gctgagactc cttgatgatg gagcaggtgc tgaatgtata aaaaatgatg gaatttactc 7020
gaggtatttt ttctcctttg ctgcaaatgt tagatatagc ttgaaagtgc atgtcaatca 7080
ctctccacgc ataagcacc cagcccactc tattccaggg agtcaatgcta tgtatgtacc 7140
aggttacaca gcaaacgcta atattcagat gaatgctcca aggaaatcag taggcagaaa 7200
tgaggaggag cgaaagtggg gctttagccg agtcagctca ggaggctcct ttctagtgct 7260
ggaggttcca cctggccccc accctgatgt gtttccacca tgcaaaatta ttgacctgga 7320
agctgtaaaa gtagaaggag aattgacct atcttggaga gccctggag aagactttga 7380
tcaggggcag gctacaagct atgaataaag aatgagtaaa agtctacaga atatccaaga 7440
tgaactttat aatgctattt tagtaataac atcaaaagca aatcctcagc aagctggcat 7500
caggagagata tttaactgtt caccccaaat ttccacgaat ggacctgaac atcagccaaa 7560
tggagaaaaca catgaaaagc acagaattta tgttgcaata cgagcaatgg ataggaactc 7620
cttacagctc gctgtatcta acattgcccc ggcgcctctg tttattccccc ccaattctga 7680
tctgtacct gccagagatt atcttatatt gaaaggagtt ttaacagcaa tgggtttgat 7740
aggaatcatt tgccttatta tagttgtgac acatcatact ttaagcagga aaaagagagc 7800
agacaagaaa gagaatggaa caaaattatt ataataaatt ctgcagatat ccatcacact 7860
ggcggcgctc cgagcaccac caccaccacc actgagatcc ggctgctaac aaagcccga 7920
aggaagctga gttggctgct gccaccgctg agcaataact agcataaacc cttggggcct 7980
ctaaacgggt cttgaggggt tttttgctga aaggaggaaac tatatccgga t 8031

```

&lt;210&gt; 255

&lt;211&gt; 401

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc feature

&lt;222&gt; 9, 67, 247, 275, 277, 397

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 255

```

gtggccagng actagaaggc gaggcgcgcg gggaccatgg cggcgccggc ggacgagcgg 60
agtcacnagg acggagaaga cgaggaagag gaggagcagt tggttctggt ggaattatca 120
ggaattattg attcagactt cctctcaaaa tgtgaaaaa aatgcaaggt tttgggcatt 180
gacactgaga ggcccattct gcaagtggac agctgtgtct ttgctgggga gtatgaagac 240
actctangga cctgtgttat atttgaagaa aatgntnaac atgctgatac agaaggcaat 300
aataaaaacg tgctaaaata taaatgccat acaatgaaga agctcagcat gacaagaact 360
ctcctgcagc agaagaagga aggagaagaa aacatanctg g 401

```

&lt;210&gt; 256

&lt;211&gt; 401

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 7, 37, 51, 79, 96, 98, 103, 104, 107, 116, 167, 181, 183,  
 194, 206, 276, 303, 307, 308, 310, 323, 332, 341, 353, 374,  
 376  
 <223> n = A,T,C or G

<400> 256  
 tgggtggnct gggatgggga accgcggtgg ctccgngga ggtttcgga ntggcatcgg 60  
 gggcggggt cgcggccgng gacggggcgg gggccnangc cgnnganctc gcggangcaa 120  
 ggccgagatg aaggagtggg tgcccgtoac caacttgggc cgcttgncca aggacatgaa 180  
 nancaagccc ctgnaggaga tctatntctt cttccctgcc ccattaagga atcaagagat 240  
 cattttgattt cttcctgggg gcctctctca aggatnaggt ttttgagat tatggcagtg 300  
 canaaannan acccctgtgc cngtccatc tncacccaac ncttccaagg gcnatttttg 360  
 tttaggcctc attncngggg ggaaccttaa cccaatttgg g 401

<210> 257  
 <211> 401  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 382, 387  
 <223> n = A,T,C or G

<400> 257  
 atgtatgtaa aaccttcoat aaaatgtaaa gggctataac aaatatgtta taaagtgtt 60  
 ctctcagccc tgaggtatatac agaatcattt gcctcagact gctgttggat tttaaaattt 120  
 ttaaaaatc tgctaagtaa tttgctatgt cttctccacc actatcaata tgctgtctc 180  
 taacaggctc cccactttct tttaatgtgc tggatgagc tttggacatg agataacctg 240  
 gcctgttcag agtgtctaca gtaagagctg gacaaactct ggaggagac agtcttttag 300  
 acagctcttt tggttgcttt ccacttttct gaaaggttca cagtaacctt ctgataata 360  
 gaaactccca gttaaagcct angctancaa ttttttttag t 401

<210> 258  
 <211> 401  
 <212> DNA  
 <213> Homo sapiens

<400> 258  
 ggagcgctag gtccgtgtac gaccgagatt aggggtgcgtg ccagctccgg gaggccgcgg 60  
 tgagggggcc ggcccaagct gccgaccoga gccgatogtc aggggtcgca ggcctcagc 120  
 tctgtggagg agcagcagta gtccgagggt gcaggatatt agaaatggct actccocagt 180  
 caatttttcat ctttgcacatc tgcattttta tgataacaga attaatctct gcctcaaaaa 240  
 gctcatatga tatcttaggt gtgccaaaat ccgcatcaga gcgccaatc aagaagacct 300  
 ttcacacagtt ggccatgaag taccaccctg acaaaaataa gaccocagat ctgaagcaaa 360  
 attcagagag attgcagaag catatgaaac actctcagat g 401

<210> 259  
 <211> 401  
 <212> DNA  
 <213> Homo sapiens

<400> 259  
 attgggtttg gaggggaggat gatgacagag gaatgcocct tggccatcac gggttttgatt 60  
 ctccagaata ttgtgggttt gatcatcaat gcagtcagat taggtcgcat tttcatgaaa 120



```

acagctcagg ctcacagaag ggcagaaact ttgattttca gccgccatgc tgtgattgcc 180
gtccgaaatg gcaagctgtg cttcatgttc cgagtgggtg acctgaggaa aagcatgac 240
attagtgcct ctgtgcgcac ccaggtggtc aagaaaacaa ctacacctga aggggaggtg 300
gttccatatt accaacttga cattcctgtt gataacccaa tcgagagcaa taacattttt 360
ctgtgggccccc ctttgatcat ctgccacgtg attgacaagc g 401

```

<210> 260

<211> 363

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 7, 9, 19, 41, 63, 73, 106, 111, 113, 116, 119, 156, 158,  
162, 187, 247, 288, 289, 290, 292, 298, 299, 300, 340

<223> n = A,T,C or G

<400> 260

```

aggaganaag gaggggggana tgaataggga tggagaggga natagtggat gagcagggca 60
canggagagg aancagaaag gagaggcaag acaggggagac acacancaca nangangana 120
caggtggggg ctgggggtggg gcatggagag cctttngant cccccaggcc acctgtctct 180
cgctggmctg ttgaaaccca ctccatggct tcttgccact gcagttgggc ccagggtctg 240
cttattnctg gaatgcaagt ggctgtggct tggagcctcc cctctggnnn anggaaannn 300
attgtccctt tatctgcttg gaatatctga gtttttccan cccggaataa aaacacacac 360
aca 363

```

<210> 261

<211> 401

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 114, 152

<223> n = A,T,C or G

<400> 261

```

cggctctccg ccgctctccc ggggtttcgg ggcacttggg tccccagtc tggctctgct 60
tcaccttccc ctgacctgag tagtcgccat ggcacaggtt ctgagaggca ctgngactga 120
cttccctgga tttgatgagc gggctgatgc anaaactctt cggaaggcta tgaaaggctt 180
gggcacagat gaggagagca tcttgactct ttgacatccc cgaagtaatg ctccagcgcca 240
ggaaatctct cgagctttta agactctgtt ttgocaggat ctctggatg acctgaaact 300
agaaactaact ggaaatttg aaaaattaat tgtggctctg atgaaacccc ctccggcttta 360
tgatgcttat gaactgaaac atgccttgaa gggagctgga a 401

```

<210> 262

<211> 401

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 7, 26, 258, 305, 358, 373, 374, 378

<223> n = A,T,C or G

<400> 262

```

agtctanaac atttctaata ttttngctt tcatatatca aaggagatta tgtgaaacta 60
tttttaataa ctgtaaagtg acatatagtt ataagatata tttctgtaca gtgagaaa 120

```

```

agttataaac atgaagaata ttgtaccatt atacattttc attctcgatc toataagaaa 180
ttcaaaagaa taatgataga ggtgaaaata tgtttacttt ctctaaatca agcctagttg 240
tcaactcaaa aattatgntg catagtttta ttttgaattt aggttttggg actacttttt 300
tccancttca atgagaaaat aaaatctaca actcaggagt tactacagaa gttctaanta 360
tttttttgct aannagcnaa aaatataaac atatgaaaat g 401

```

<210> 263

<211> 401

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 232, 290, 304, 326, 383

<223> n = A,T,C or G

<400> 263

```

ctgtccgacc aagagaggcc ggccgagccc gagccttggg cttttgcttt ctggcgaggg 60
gatctcgccc ggttttaggag gcggcgctga tctctggagg aagaggcagc tacggcgggc 120
gcggcggttg cggctagggg ggccgccaat aaaggggccc ccgccgggtg atgcgggtgac 180
cactcgggca ggcccaggag ctgagtgggc ccggcccttc agcccgctcc gncggaccgg 240
ctttctctaa ctctccatct tctctctccc accgagatcg ccgagggcgg ctcaggctcc 300
ctanccctct ccccgctcct tccccncccc cgtccccgcc ccggggggcg ccgccaccgg 360
ctccccacca tggctctgaa ganaatccac aaggaattga a 401

```

<210> 264

<211> 401

<212> DNA

<213> Homo sapiens

<400> 264

```

aacaccagcc actccaggac cctgaaggc ctctaccagg tcaccagtgt tctgcgccta 60
aagccacccc ctggcagaaa ctccagctgt gtgttttgga atactcacgt gagggaaatt 120
actttggcca gcattgacct tcaaaagtcag atggaaccca ggaccatccc aacttgggtg 180
cttcacattt tcatcccttc ctgcatcatt gctttcattt tcatagccac agtgatagcc 240
ctaagaaaaa aactctgtca aaagctgtat tcttcaaaa acacaacaaa aagacctgtc 300
accacaacaa agaggggaagt gaacagtgct gtgaattctga acctgtgtgc ttgggagcca 360
gggtgacctg atatgacatc taaagaagct tctggactct g 401

```

<210> 265

<211> 271

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 59

<223> n = A,T,C or G

<400> 265

```

gccacttctt gtggacatgg gcagagcgct gctgccagtt cctggtagcc ttgaccacna 60
cgctgggggg tcttttgtat ggtcatgggt ctcatattgca ctctgggggtg tgggattcaa 120
gttagaagtt tctagatctg gccgggcgca gtggctcaca cctgtaatcc cagcacttta 180
ggagctcgag gcaggcggat catgagtgca ggagatcgag accgtctctgg ctaacacagt 240
gaaaccccgct ctctactaaa aatacaaaaa a 271

```

<210> 266

<211> 401

<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<222> 45  
<223> n = A,T,C or G

<400> 266  
attcataaat ttagctgaaa gatactgatt caatttgat acagngaata taaatgagac 60  
gacagcaaaa ttttcatgaa atgtaaaata tttttatagt ttgttcatac tatatgaggt 120  
tctatttttaa atgactttct ggatttttaa aaatttcttt aaatacaatc atttttgtaa 180  
tatttttttt atgcttatga tctagataat tgcagaatat cattttatct gactctgtct 240  
tcataagaga gctgtggcgc aattttgaac atctgttata gggagtgtac aaattagaag 300  
gcaatgtgga aaaacaattc tgggaaagat ttctttatat gaagtccctg ccactagcca 360  
gcaatcctaa ttgatgaaag ttatctgttc acaggcctgc a 401

<210> 267  
<211> 401  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<222> 116, 247, 277, 296, 307, 313, 322, 323, 336, 342, 355, 365,  
377, 378, 397  
<223> n = A,T,C or G

<400> 267  
gaagaggcat cacctgatcc cggagacctt tggagttaag aggcggcgga agcgagggcc 60  
tgtggagtcg gatcctcttc ggggtgagcc agggctcgcg cgcgcgctg tctcanaact 120  
catgcagctg ttcccgcgag gcctgtttga ggacgcgctg ccgcccatcg tgctgaggag 180  
ccaggtgtac agccttctgc ctgacaggac cgtggcgcac cggcagctga aggagcttca 240  
agagcanggg gagacaaaat cgtccagctg ggcttncact tggatgccca tggaanttat 300  
tctttcnctt ganggactta cnngggaccc aagaancctt tncaaggggc ccttngtgga 360  
tggncccca aaccccnnta tttgcccttg ggggggncca a 401

<210> 268  
<211> 223  
<212> DNA  
<213> Homo sapiens

<400> 268  
tcgccatggt ggccagcgctg gtcttgaact cctgacttta agtgatccac ccgcctcaac 60  
ctcccaaggt gctgggatta caggtgtgag ccacgcgcgc tggcctgata catactttta 120  
gaatcaagta gtcaagcact ttttctgttc atttttctaa aaagtaaata tacaatgtt 180  
ttgttttttg ttttttttgt ttgtttgttt ctgttttttt ttt 223

<210> 269  
<211> 401  
<212> DNA  
<213> Homo sapiens

<400> 269  
actatgtaaa ccacattgta ctttttttta ctttggcaac aaatatattat acatacaaga 60  
tgctagttca ttgtgaattt tctcccaact tatccaagga tctccagctc taacaaaatg 120  
gtttattttt atttaaatgt caatagtgtt tttttaaaat ccaaatcaga ggtgcaggcc 180  
accagttaaa tgccgtctat caggtttttg gccttaagag actacagagt caaagctcat 240

```

ttttaaaagga gtaggacaaa gttgtcacag gtttttgttg ttgtttttat tgcccccata 300
attacatggtt aatttcattt tatatcaggg attctatatta cttgaagact gtgaagtgtc 360
cattttgtct cattgttttc tttagacataa ctaggatcca t 401

```

<210> 270

<211> 401

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 240, 382

<223> n = A,T,C or G

<400> 270

```

tggtgtttga ttcaactcag caactgcttg tatctgcacc ctacctctct ttagaggctg 60
ccttgtaaac tgaataatgc acctgacttc gagcaagact ctttccttag gttctggatc 120
tgtttgagcc ccattggcact gagctggaat ctgagggtct tgttccaagg atgtgatgat 180
gtggggagaat gttctttgaa agagcagaaa tccagtctgc atggaacacg cctgtagagn 240
agaagtttcc agtgataagt gttcactgtt ctaaggaggt acaccacagc tacctgaatt 300
ttcccaaat gagtgcttct gtgcgttaca actggccttt gtacttgact gtgatgactt 360
tgttttttct tttcaattct anataaacat gggaaaaaat g 401

```

<210> 271

<211> 329

<212> DNA

<213> Homo sapiens

<400> 271

```

ccacagcctc caagtcaggt ggggtggagt cccagagctg cacagggttt ggcccaagtt 60
tctaaggagg gcacttcttc ccctcgccca taagtgccag ccctgctgg ctggtgctgt 120
agcccccacg acagccccc gcccgcaggg cctgccttct cagggaacttc tggcgggctc 180
gaggcaagcg atggagtgag acccaggagc cggacacttc tcaggaaagt gcttttccca 240
acccccagcc cccaccgggt ggttcttctt gttctgtgac tgtgtatagt gccaccacag 300
cttatggcat ctcatggagg acaaaaaaa 329

```

<210> 272

<211> 401

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 1, 7, 12, 21, 61, 62, 66, 72, 78, 88, 90, 92, 98, 117, 119, 128, 130, 134, 142, 144, 151, 159, 162, 164, 168, 169, 177, 184, 185, 188, 194, 202, 204, 209, 213, 218, 223, 231, 260, 272, 299, 300, 306, 321, 322, 323, 331, 335, 336, 338

<223> n = A,T,C or G

<221> misc\_feature

<222> 341, 342, 343, 345, 346, 351, 358, 360, 362, 363, 387, 390, 392

<223> n = A,T,C or G

<400> 272

```

nggctgntaa cntcggaggt nacttctctg actatcctgg agaccccttc cgcttccacg 60
nncatnatat cntcatngc tgggccntn angacacnat cccactccaa cacctgmgng 120
atgctggncn cctnggaacc ancntcagaa ngacctgnt cntntgtntt ccgcaanctg 180

```

```

aagannaangc gggntacacc tncntgcant ggnccacnct gcngggaact ntacacacct 240
acgggatgtgt gctgcgcgan gagccaagag cntttctgga tgattcccca gcctcttgnn 300
agggantcta caacattgtt nnttaccttt ntcmnncgc nmntnntgga ntacaggngn 360
tnntaacactt acatcttttt tactgcncn tnccttggtgg g 401

```

&lt;210&gt; 273

&lt;211&gt; 401

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 399

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 273

```

cagcaccatg aagatcaaga tcctgcgacc ccagagcgc aagtactcgg tgtggatcgg 60
tggctccatc ctggcctcac tgtccacctt ccagcagatg agcaggagta 120
cgagcagctg gggccctcca tcgtccaccg caaatgcttc taaacggact cagcagatgc 180
gtacgatttg ctgcatgggt taattgagaa tagaaatttg ccctggcaa atgcacacac 240
ctcatgttag cctcacgaaa ctggaataag ccttcgaaaa gaaattgtcc ttgaagcttg 300
tatctgatat cagcactgga ttgtagaact tgttgctgat ttgaccttg tattgaagtt 360
aactgttccc cttggtatta acgtgtcagg gctgagtnt c 401

```

&lt;210&gt; 274

&lt;211&gt; 401

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 274

```

ccaccacac ccacgcgcgc ctctgtcgcc tcttctccgg gagccagtcc gcgccaccgc 60
cgccgcgcc gccatgcga cctcgcgcag ccatgtccac caggtccgtg tctctgtcct 120
cctacgcgag gatgttcggc ggcccgggca ccgcgagcgc gccgagctcc agccggagct 180
acgtgactac gtcacaccgc acctacagcc tgggcgcgcg gctgcgcgcc agcaccagcc 240
gcagcctcta cgctctgtcc cggggcggcg tgtatgccac gcgctcctct gcctgcgcc 300
tgcgagcag cgtgcccggy gtgcggtcc tcgagactc ggtggacttc tcgtggccg 360
acgccatcaa caccgagttc aagaacacc gccaccaaga g 401

```

&lt;210&gt; 275

&lt;211&gt; 401

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 275

```

ccacttccac cactttgttg agcagtgcct tcagcgcaac ccgatgccca ggtatccctg 60
ctggcctggg cctgggcttc gggagagcag aggggtgcct gaagggttaag gccagggtgt 120
gaagggaactt acctcccaaa ggttctgcag gggaatctgg agctacacac aggagggatc 180
agctcctggg tgtgtcagag gccagcctgg ggagctctgg ccaactgtct ccatgagctg 240
aggagagggy agaggggacc cgaggctgag gcataagtgg caggatttcg ggaagctggg 300
gacacggcag tgatgctgcg gtctctcttc ccctttccct ccaggccag tgccagcacc 360
ctcctgaacc actctttctt caagcagatc aagcgacgtg c 401

```

&lt;210&gt; 276

&lt;211&gt; 401

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

```

<221> misc_feature
<222> 11
<223> n = A,T,C or G

<400> 276
tctgatattg ntaccottga gccacctaag ttagaagaaa ttggaaatca agaagttgtc 60
attgtttgaag aagcacagag ttcagaagac tttaacatgg gctcttcctc tagcagccag 120
tatactttct gtcagccaga aactgtattt tcatctcagc ctagtgtatga tgaatcaagt 180
agtgatgaaa ccagtaataca gcccagtcct gcctttagac gacgcogtgc taggaagaag 240
accgttttctg cttcagaato tgaagaccgg ctagtgtgtg aacaagaaac tgaaccttct 300
aaggagttga gtaaacgtca gttcagtagt ggtctcaata agtgtgttat acttgctttg 360
gtgattgcaa tcagcatggg atttggccat ttctatggca c 401

<210> 277
<211> 401
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 227, 333
<223> n = A,T,C or G

<400> 277
aactttggca acatatctca gcaaaaaacta cagctatgtt attcatgcc aataaaaagc 60
tgtgcagagg agtggctgca atgaggtcac aacggtgtgt gatgtaaaag agatcttcaa 120
gtctcatca cccatccctc gaactcaagt cccgctcatt acaaatctct cttgccagtg 180
tccacacatc ctgcccctac aagatgttct catcatgtgt tacgagnggc gctcaaggat 240
gatgcttctt gaaaattgct tagttgaaaa atggagagat cagcttagta aaagatccat 300
acagtgggaa gagaggctgc aggaacacgc ganaacagtt caggacaaga agaaaaacgc 360
cgggcgcacc agtcgtagta atccccccaa accaaaggga a 401

<210> 278
<211> 401
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 322, 354
<223> n = A,T,C or G

<400> 278
aatgagtggt agaccacaaa tgaatgccgg gaggatgaaa tgtgttggaa ttatcatggc 60
ggcttcogtt gttatccaag aaatccttgt caagatccct acattctaac accagagaac 120
cgatgtgttt gcccagttct aaatgccatg tgcccgagac tgcccagtc aatagtctac 180
aaatacatga gcatccgato tgataggtct gtgccatcag acatcttcca gatacaggcc 240
acaactattt atgccaacac catcaatact ttccggatta aatctggaaa tgaanaatgga 300
gagctctact acgacaacaa anccctgtaa gtgcaatgct tgtgtcgtg aagncattat 360
caggaccaag agaacatatc gtggacctgg agatgtctgac a 401

<210> 279
<211> 401
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

```

&lt;222&gt; 30, 35, 81, 88, 180, 212, 378, 384, 391

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 279

```

aaattattgc ctctgatata tacctaagtn aacanaacat taatacctaa gtaaacataa 60
cattacttgg aggtgttcag ntcttaantg aaactgtatt tgaactttt aagtatactt 120
taggaaccaa gcatgaacgg cagctctaga taccagaacac atctactttg gtagcttggn 180
gccattatcc tgtggaatct gatagtctgt gnaagcatgt attgatggga ctgaagaca 240
tctttggaaa tgatgagatt atttctctgt ttaaaaaaaa aaaaaatctt aaattctctac 300
aatgtgaac tgaaactaat aattttgato ctgatgtatg ggacagcgta tctgtaccag 360
gctctaaata acaaaagnta gggngacaag nacatgttcc t 401

```

&lt;210&gt; 280

&lt;211&gt; 326

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 280

```

gaagtggaaat tgtataatcc aattcgataa ttgatctcat gggctttccc tggaggaaag 60
gtttttttttg ttgttttttt ttttaagaact tgaactttgt aaactgagat gtctgtagct 120
tttttgccca tctgtagtgt atgtgaagat ttcaaaacct gagagcactt tttctttgtt 180
tagaattatg agaaaggcac tagatgactt taggatttgc atttttccct ttattgcctc 240
attttctgtg acgccttgtt ggggagggaa atctgtttat tttttctctac aaataaaaaa 300
ctaagattct atatcgcaaa aaaaaa 326

```

&lt;210&gt; 281

&lt;211&gt; 374

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 281

```

caacgcgttt gcaaataatc cccctggtagc ctacttccct acccccgaat attgtaaga 60
tcgagcaatg gtttcaggac atgggtttct tttctctgtg atcattcaag tgcctactgc 120
atgaagaactg gcttgtctca gtgtttcaac ctacaccagg ctgtctcttg gtccacacct 180
cgctccctgt tagtgcgcta tgacagcccc catcaaatga ccttggccaa gtccaggttt 240
ctctgtgtgc aaggttggtt ggctgattgg tggaaagtag ggtggaccaa aggaggccac 300
gtgagcagtc agcaccagtt ctgcaccagc agcgctccg tctagtggg tgttctgtt 360
tctctgtgcc ctgg 374

```

&lt;210&gt; 282

&lt;211&gt; 404

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 26, 27, 51, 137, 180, 222

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 282

```

agtgtgttgg aattcccgca tctctanncc cgactcacac aaggcagagt ngccatggag 60
aaaattccag tgtcagcatt cttgtctctt gtggccctct cctacactct ggccagagat 120
accacagtcg aacctgnagc caaaaaggac acaaaaggact ctgcacccaa actgcccacc 180
acctctctca gaggttgggg tgaccaactc atctggactc anacatatga agaagctcta 240
tataaatcca agacaagcaa caaacctctg atgattatct atcaattgga tgagtgcaca 300
cacagtcagc ctttaagaaa agtgtttgct gaaaataaag aaatccagaa attggcagag 360
cagtttgtcc tctcaatct ggtttatgaa acaactgaca aaca 404

```

<210> 283  
 <211> 184  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 26  
 <223> n = A,T,C or G

<400> 283  
 agtgtggtgg aattcacttg cttaanttgt gggcaaaaga gaaaagaag gattgatcag 60  
 agcattgtgc aatacagttt cattaactcc ttccctcgct ccccaaaaaa ttgtaatttt 120  
 tttttcaaca ctcttacacc tgttatggaa aatgtcaacc ttgtaagaa aaccaaata 180  
 aaaa 184

<210> 284  
 <211> 421  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 147, 149  
 <223> n = A,T,C or G

<400> 284  
 ctattaatcc tgccacaata tttttaatta cgtacaaaga tctgacatgt caccaggga 60  
 cccatttcac ccaactgctct gtttgccgc cagtcttttg tctctctctt cagcaatggt 120  
 gaggcgata ccctttctctc gggaanana aatccatggt ttgttgccct tgccaataac 180  
 aaaaatgttg gaaagtcgag tggcaagct gttgccattg gcattcttca cgtgaaccac 240  
 gtcaaaagat ccagggtgac tctctctggt ggtgatcaca ccaattcttc ctaggttagc 300  
 acctccagtc accatacaca ggttaccagt gtcgaacttg atgaaatcag taactctgcc 360  
 agtctctaaa tcaatctgaa tggatatcatt caccttgatg aggggatcgg ggtagcggat 420  
 g 421

<210> 285  
 <211> 361  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 34, 188  
 <223> n = A,T,C or G

<400> 285  
 ctgggtggta actctttatt tcattgtccg gaanaaagat gggagtggga acagggtgga 60  
 cactgtgcag gcttcagctt ccaactcggg caggattcag gctatctggg accgcaggga 120  
 ctgccagggt cacagcctg gctcccaggg caggcaggca aggtgacggg actggaagcc 180  
 cttttcanag ccttggagga gctggtccgt ccacaagcaa tgagtgcac tctgcagttt 240  
 gcaggggatg gataaacagg gaaacactgt gcattcctca cagccaacag tgtagggttt 300  
 ggtgaagccc cggcgctgag ctaagctcag gctgttcacg ggagccacga aactgcagggt 360  
 a 361

<210> 286  
 <211> 336  
 <212> DNA



<213> Homo sapiens

<220>

<221> misc\_feature

<222> 40, 68, 75, 127, 262

<223> n = A,T,C or G

<400> 286

```

tttgagtggc agcgccttta tttgtggggg ccttcaaggn agggctcgtg ggggcagcgg 60
ggaggaaanag coganaaact gtgtgaccgg ggccctcaggt ggtgggcatt gggggctcct 120
cttgcanatg cccattggca tcaccgggtgc agccattggt gcagcgggt accggtcctt 180
tcttgttcaa catagggtag gtggcagcca cgggtccaac tcgcttgagg ctggggccctg 240
ggcgctccat tttgtgttcc angagcatgt ggttctgtgg cgggagcccc acgcaggccc 300
tgaggatggt ctgcgtgcag ctgcgctggc ggaaaa 336

```

<210> 287

<211> 301

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 15, 33, 44, 53, 76, 83, 107, 117, 154, 166, 192, 194, 207, 215, 241, 246

<223> n = A,T,C or G

<400> 287

```

tgggtaccaa atttntttat ttgaaggaat ggnacaaatc aaanaactta agnggatggt 60
ttggtacaac ttatanaaaa ggnaaaggaa accccaacat gcatgcncctg ccttggngac 120
cagggaagtc accccaaggc tatggggaaa ttancccgag gcttancttt cattatcact 180
gtctccacgg gmgngcttgt caaaaanata ttccnccaag ccaaattcgg gcgctcccat 240
nttgcncag ttggtcacgt ggtcacccaa ttctttgatg gctttcaact gctcattcag 300
g 301

```

<210> 288

<211> 358

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 39, 143, 226

<223> n = A,T,C or G

<400> 288

```

aagtttttaa acttttttatt tgcataattaa aaaaattgng cattccaata attaaaaatca 60
tttgaaacaaa aaaaaaaatg gcaactctgat taaactgcat tacagcctgc aggacacctt 120
ggggcagcctt ggttttactc tanatttcaac tgtgtccca cccactctct tccacccac 180
ttcttccttc accaacatgc aagttcttct ctccctgcc agccanatag atagacagat 240
gggaaaggca ggcggggcct togttgtcag tagttctttg atgtgaaagg ggcagcacag 300
tcatttaaac ttgatccaac ctctttgcat cttacaaagt taaacagcta aaagaagt 358

```

<210> 289

<211> 462

<212> DNA

<213> Homo sapiens

<220>

<221> misc feature  
 <222> 87, 141, 182, 220, 269, 327  
 <223> n = A,T,C or G

<400> 289  
 ggcatcagaa atgctgttta tttctctgct gctcccaagc tggctggcct ttgcagagga 60  
 cgagacaaca gatgcatagt tgggganaaa gggaggacag gttccaggat agaggggtgca 120  
 ggctgaggga ggaagggttaa naggaaggaa ggccatcctg gatcccccaca tttcagctctc 180  
 anatgaggac aaagggactc ccaagccccc aaatcatcan aaaacaccaa ggagcaggag 240  
 gagcttgagc aggcoccagc gagcctcana gccataccag ccactgtcta ctcccatcc 300  
 tcctctccca ttccctgtct gcttcanacc acctcccagc taagcccccag ctccattccc 360  
 ccaatcctcg cccttgccag cttgacagtc acagtgcctg gaattccacc actgaggctt 420  
 ctcccagttg gatttagacg tcgccctgtt agcatgctgc cc 462

<210> 290  
 <211> 481  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc feature  
 <222> 44, 57, 122, 158, 304, 325, 352, 405  
 <223> n = A,T,C or G

<400> 290  
 tactttccta aactttatta aagaaaaaag caataagcaa tggnggttaa tctctanaac 60  
 ataccocaatt ttctgggctt cctcccccga gaatgtgaca ttttgatttc caaacatgcc 120  
 anaagtgtat ggttcccaac tgtactaaag taggtganaa gctgaagtc tcaagtgttc 180  
 atcttccaac ttttccagct ctgtggtctg tctttggatc agcaataatt gccatgaacag 240  
 ctactatggc ttocgttgatt ttgtctgta gctctctgag ctccctctatg tgcagcaact 300  
 gcanaatttg agcagcttca ttaanaactg catctcctgt gtcaaaaacca anaatatgtt 360  
 tgtctaaagc aacaggtaag cctctctttg ttgtatttgc cttanacaact gcatcctgtg 420  
 tcaggcgctc ctgaacccaa atccgaattg ccttaagcat taccaggtaa tcatcatgac 480  
 g 481

<210> 291  
 <211> 381  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc feature  
 <222> 79, 166, 187, 208, 219, 315  
 <223> n = A,T,C or G

<400> 291  
 tcatagtaat gtaaaaccat ttgtttaatt ctaaatcaaa tcactttcac aacagtga 60  
 attagtgaact ggtaaggng tgccactgta catatcatca ttttctgact ggggtcagga 120  
 cctggtccta gtcacaaagg gtggcaggag gaggggtggag gctaanaaca cagaaaacac 180  
 acaaaaanaa ggaagctgc cttggcanaa ggatgaggng gtgagcttgc cgaaggatgg 240  
 tgggaagggg gctccctgtt ggggccgagc caggagctcc aagtgcagctc tcctgcctta 300  
 cttagctcct ggcanaaggt gagtggggac ctacgaggtt caaaatcaaa tggcatttgg 360  
 ccagcctggc tttactaaca g 381

<210> 292  
 <211> 371  
 <212> DNA  
 <213> Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 32, 55, 72, 151, 189, 292

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 292

```

gaaaaaataa tccgtttaat tgaaaaaacct gnaggatact attccactcc cccanattgag 60
gaggctgagg anaccaaacc cctacatcac ctctgtagcca ctctgtgatac tcttcacgag 120
gcagcaggca aagacaattc ccaaaaacctc nacaaaagca attccaagggt ctgtgtgcagc 180
taccaccanc acatttttcc tcagccagcc cccaatcttc tccacacagc cctccttatg 240
gatcgcttc tcgttgaaat taatcccaca gccacagta acattaatgc ancaggagtc 300
ggggactcgg ttcttcgaca tggaagggat tttctcccaa tctgtgtagt tagcagcccc 360
acagcactta a                                     371

```

&lt;210&gt; 293

&lt;211&gt; 361

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 75, 196, 222

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 293

```

gatttaaaag aaaacacttt attgttcagc aattaaaaat tagccaaata tgtatttttc 60
tcataaatt attngatgt tatcaacatc aagtaaaatg ctcatatttc tcatattgctt 120
ctgttcattgt tttcttgaaac acgtcttcaa tttctcttc aaaatgtgtc atgcacacact 180
tgaggtaacg aagcanaagt atttttaaac atgaacgcta anaacattca tctacagcaa 240
cctatatgct caatacatgc cgcgtgatcc tagtagtttt ttccacaacct tctacaagtt 300
tttggaanaac atctgttatg atgactttca tacaccttca cctcaaaaggc tttctgtcac 360
c                                     361

```

&lt;210&gt; 294

&lt;211&gt; 391

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 26, 77, 96, 150, 203, 252, 254, 264, 276

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 294

```

tattttaaag ttttaattat attcanaaaa aatcgagcga ataactttct ctgaaaaaat 60
atattgactc tgtatanacc acagttattg gggganaagg gctggtaggt taaattatcc 120
tattttttat tctgaaaatg atattaatan aaagtccctg ttccagtcgt attataaaga 180
tacatatgcc caaaatggct ganaataaat acaacaggaa atgcaaaagc tgtaaagcta 240
agggcatgca ananaaaatc tcanaataacc caaagnggca acagggaacg tttggctgga 300
atttgaagtt atttcagtc tctttgtctt tggctccatg tttcaggatg cgtgtgaaat 360
cgatgtaatt gaaattcccc tttttatcaa t                                     391

```

&lt;210&gt; 295

&lt;211&gt; 343

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 145, 174, 205, 232  
 <223> n = A,T,C or G

<400> 295  
 ttctttttgtt ttattgataa cagaaactgt gcataattac agatttgatg aggaatctgc 60  
 aaataataaa gaatgtgtct actgccagca aaatacaatt attccatgcc ctctcaacat 120  
 acaaatatag agttcttcac accanattggc tctggtgttaa caaagccatt ttanatgttt 180  
 aattgtgctt ctacaaaacc ttcanagcat gaggtagttt cttttaccta cnatattttc 240  
 cacattttcca ttattacact tttagtgagc taaaatcctt ttaacatagc ctgcggatga 300  
 tctttcacaa aagccaagcc tcattttacaa agggttttatt tct 343

<210> 296  
 <211> 241  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 96, 98, 106, 185  
 <223> n = A,T,C or G

<400> 296  
 ttcttgataa ttggtgtgtt ttgtgaaaaa gtttttgttt ttcttctcag tcaactgaat 60  
 taatttctcta ctttgccctc ctgatgccca catgananaa ottaanataa tttctaacag 120  
 cttccacttt ggaaaaaaa aaaacctgtt ttctcatgg aaacccagga gttgaaagtg 180  
 gatanatgcg tctcaaaatc taaggctctg ttcagcttta cattatgtta cctgacgttt 240  
 t 241

<210> 297  
 <211> 391  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 12, 130  
 <223> n = A,T,C or G

<400> 297  
 gttgtggctg anaatgctgg agatgctcag ttctctccct cacaaggtag gccacaaatt 60  
 ctgtgtgggt ccttcacatc tggggtcttc aggcaccagc catgcctgcc gaggagtgtc 120  
 gtcaggacan accatgtccg tgctaggccc aggcacagcc caacctctcc tcatccaagt 180  
 ctctccagg tttctggtcc cgatggggca ggatgacccc tccagtggct ggtaccocac 240  
 catccacta cccctcacat gctctcactc tccatcaggt ccccaatcct ggcttccctc 300  
 ttcaagcaact ctcaaaagaa aggaaggata aaacctaaat aaaccagaca gaagcagctc 360  
 tggaaaagta caaaaagaca gccagaggtg t 391

<210> 298  
 <211> 321  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 14, 30, 76, 116, 201, 288, 301  
 <223> n = A,T,C or G

```

<400> 298
caagccaaac tgnntccagc tttattaaan atactttcca taaacaatca tggattttca 60
ggcaggacat gggcanacaa tegttaacag tatacaacaa ccttntttca 120
atggactacc aaaaatcaaa aagccactat aaaacccaat gaagtcttca tctgatgctc 180
tgaacaggga aagtttaaag ngagggttga catttcacat ttagcatgtt gttaacaac 240
ttttcacaag ccgacctga ctttcaggaa gtgaaatgaa aatggcanaa tttatctgaa 300
natccacaat ctaaaaatgg a 321

```

```

<210> 299
<211> 401
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 104, 268, 347
<223> n = A,T,C or G

```

```

<400> 299
tatcataaag agtgttgaag tttatttatt atagcaccat tgagacattt tgaatttga 60
atttgtaaaa aaataaaaca aaaagcattt gaattgtatt tggnggaaca gcaaaaaaag 120
agaagtatca tttttctttg tcaaatata ctgtttccaa acattttgga aataaataac 180
tggaattttg tgggtcactt gcaactggtg acaagattag aacaaggaga acacatatgg 240
agttaaattt tttttgttgg gatttcanat agagtttggt ttataaaaag caaacagggc 300
caacgtccac accaaattct tgaatcaggac caccaatgtc atagggngca atatctacaa 360
taggtagtct cacagccttg cgtgttcgat attcaaaagc t 401

```

```

<210> 300
<211> 188
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> 48
<223> n = A,T,C or G

```

```

<400> 300
tgaatgcttt gtcatattaa gaaagttaaa gtgcaataat gtttgaanac aataagtgtt 60
gggtgatctt gtttctaata agataaaact ttttgtctt gctttatctt attaggagg 120
tgtatgtcag tgtataaaac atactgtgtg gtataacagg ctaataaat tctttaaaag 180
gaaaaaaa 188

```

```

<210> 301
<211> 291
<212> DNA
<213> Homo sapiens

```

```

<400> 301
aagattttgt tttattttat tatggctaga aagacactgt tatagocaaa atcgccaatg 60
acactaaaga aatcctctgt gcttttcaat atgcaaatat atttcttcca agagttgccc 120
tgggtgact tcaagagttc atgttaactt cttttctgga aacttctctt tcttagttgt 180
tgtattcttg aagagcctgg gccatgaaga gcttgcttaa gttttgggca gtgaactctt 240
tgatgttctg gcagtaagtg tttatctgga ctgcaatgag cagcaggtcc a 291

```

```

<210> 302
<211> 341

```

<212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 25  
 <223> n = A,T,C or G

<400> 302  
 tgatttttca taatttttatt aaatnatcac tgggaaaactaatggttcgc gtatcacaca 60  
 attacactac aatctgatat gagtggtaaa accagccaat ggaatccagg taaagtacaa 120  
 aaacgccacc ttttattgtc ctgtottatt tctcggaagg gagggttcta ctttacacat 180  
 ttcatgagcc agcagtggac ttgagttaca atgtgtaggt tccttgtggt tatagctgca 240  
 gaagaagcca tcaaattctt gaggacttga catctctcgg aaagaagcaa actagtggat 300  
 ccccgggct gcaggaattc gatatcaagc ttatcgatac c 341

<210> 303  
 <211> 361  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 15, 27, 92, 124, 127, 183, 198, 244, 320  
 <223> n = A,T,C or G

<400> 303  
 tgcagacagt aaatnaattt tatttngtt cacagaacat actaggcgat ctgcagcgtc 60  
 gctccgtgac agcccaccaa cccccaacco tntacctcgc agccacccta aaggcgactt 120  
 caanaanagc gaaggatctc acggatctca ttctaatgg tccgcogaag tctcacacag 180  
 tanacagacg gagttganat gctggaggat gcagtcacat cctaaactta cgaccaccca 240  
 ccanacttca tcccagccgg gacgtctctc cccaccggag tctctcccat ttcttctctc 300  
 accttgcgcg agttccaggn gtctgtcttc caccagtcctc acaaagctca ataaatacca 360  
 a 361

<210> 304  
 <211> 301  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 23, 104, 192  
 <223> n = A,T,C or G

<400> 304  
 ctctttacaa cagcctttat ttncggccct tgatcctgct cggatgctgg tggaggccct 60  
 tagctccgac cgccaggctc tgtgccgctc cccgcaggc gcanattcat gaacacgggtg 120  
 ctacggggct tgaggccgta ctccccagc gggagctggt cctccagggg cttccctcgt 180  
 aaggtcagc anaacaggtc gtccctgcaca cctccagcc cgctcaattg ctgcttcagg 240  
 tgggccacgg tctgcgtcag ccgcacctcg taggtgctgc tgcggccctt gttattcttc 300  
 a 301

<210> 305  
 <211> 331  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 3, 36, 60, 193, 223  
 <223> n = A,T,C or G

<400> 305  
 ganaggctag taacatcagt tttattgggt tggggnggca accatagcct ggctgggggn 60  
 ggggctggcc ctcacaggtt gttgagttcc agcagggtct ggtccaaggt ctgggtgaatc 120  
 tcgacgtttct cctccttggc actggccaaag gtctcttcta ggtcatcgat ggttttctcc 180  
 aactttggca canacctctc ggcaaaactc gtctgggtct cancctcctt cagcttctctc 240  
 tccaacagtt tgatctctc ttcataatcta tcttctttgg gggaataactc ctctctctgag 300  
 gccatcaggg acttgagggc ctggtccatg g 331

<210> 306  
 <211> 457  
 <212> DNA  
 <213> Homo sapiens

<400> 306  
 aatatgtaaa ggtaataact tttattatat taaagacaat gcaaacgaaa aacagaattg 60  
 agcagtgcaa aatttaaagg actgttttgt tctcaaagtt gcaagtttca aagccaaaag 120  
 aattatatgt atcaaatata taagtaaaaa aaagttagac ttccaagcct gtaatcccg 180  
 cactttggga ggtcgaggca ggtggatcac taacattaaa aagacaacat tagattttgt 240  
 cgatttatag caattttata aatatataac ttgtcactt ggatcctgaa gcaaaaaaat 300  
 aaagtgaatt tgggattttt gtacttggtta aaaagtttaa caccctaaat tcacaactag 360  
 tggatcccc gggctgcagg aattcgatat caagcttatc gataccgtcg acctcgaggg 420  
 ggggcccggt acccaattcg ccctatagtg agtcgta 457

<210> 307  
 <211> 491  
 <212> DNA  
 <213> Homo sapiens

<400> 307  
 gtgcttggac ggaacccggc gctcgttccc cccccggcc gccgcgccat agccagccct 60  
 ccgtcacctc ttaccgcgac cctcggagtg ccccaaggcc ccgcgcgcgc ctccagcgcc 120  
 gcgcagccac cgccgcgcgc gccgcctctc cttagtcgcc gccatgacga ccgcgtccac 180  
 ctccagcagg cgccagaact accaccagga ctcagaggcc gccatcaacc gccagatcaa 240  
 cctggagctc tacgcctcct acgtttacct gtccatgtct tactactttg accgcgatga 300  
 tgtggtcttg aagaactttg ccaaataact tcttcaccaa tctcatgagg agaggggaaca 360  
 tgctgagaaa ctgatgaagc tgcagaacca acgaggtggc cgaatctctcc ttcaggatat 420  
 caagaaacca gactgtgatg actgggagag cgggctgaat gcaatggagt gtgcattaca 480  
 tttggaaaaa a 491

<210> 308  
 <211> 421  
 <212> DNA  
 <213> Homo sapiens

<400> 308  
 ctacagcgtt cttctttctt ggtttgatcc tgactgctgt catggcgtgc cctctggaga 60  
 agggccctgga tgtgatggtg tccaccttcc acaagtactc gggcaaaagag ggtgacaagt 120  
 tcaagctcaa caagtcagaa ctaaaaggagc tgctgaccgc ggagctgcgc agcttcttgg 180  
 ggaaaaaggac agatgaagct gctttccaga agctgatgag caacttggac agcaacaggg 240  
 acaacagaggt ggacttccaa gagtactgtg tcttctctgc ctgcatcgcc atgatgtgta 300  
 acgaattctt tgaaggcttc ccagataaag agccccaggaa gaaatgaaaa ctctctgat 360  
 gtgggtgggg ggtctgccag ctggggccct ccctgtcgcc agtggggcact ttttttttcc 420  
 c 421

&lt;210&gt; 309

&lt;211&gt; 321

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 309

```

accaaattggc ggatgacgcc ggtgcagcgg gggggcccg gggccctggt ggccttggga 60
tggggaaccg cggtggcttc cgcggagggt tcggcagtg catccggggc cggggctcgg 120
gcggtggacg cggccggggc cgaggccgcg gagctcgcg aggcaaggcc gaggataagg 180
agtggatgcc cgtaccaaac ttggggccgct tggtaacga catgaagatc aagtcctcgg 240
aggagatcta tctcttctcc ctgccattta aggaatcaga gatcattgat ttcttctcgg 300
gggcctctct caaggatgag g

```

&lt;210&gt; 310

&lt;211&gt; 381

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 310

```

ttaaccagcc atattggctc aataaatagc ttcggtaagg agttaatttc cttctagaaa 60
tcagtgccta ttttctctgg aaactcaatt ttaaatagtc caattccatc tgaagccaag 120
ctgttgctcat ttctattcgg tgacattctc tcccatgaca ccagaagggg gcagaagaa 180
cacatttttc atttatagat gtttgcatcc tttgtattaa aattattttg aagggggtgc 240
ctcattggat ggcttttttt ttttctctcc agggagaagg ggagaaatgt acttggaaat 300
taatgtatgt ttacatctct ttgcaaatcc ctgtacatag agatatattt ttttaagtgt 360
aatgtaacaa catactgtga a

```

&lt;210&gt; 311

&lt;211&gt; 538

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 311

```

tttgaattta caccaagaac ttctcaataa aagaaaatca tgaatgctcc acaatttcaa 60
cataaccaca gagaagttta ttcttaaca ttgtgttcta tgattatttg taagaccttc 120
accaagtctc gatattcttt aaagacatag ttaaaaattg cttttgaaaa tctgtattct 180
tgaaaatata cttgtgtgtg attaggtttt taaaataccag ctaaaaggatt acctcaactg 240
gtcatcagta cctcctattt cagctcccca agatgatgtg tttttgctta ccttaagaga 300
ggttttcttc ttatttttag ataattcaag tgcttagata aattatgttt tctttaagtgt 360
tttatggtaa actcttttaa agaaaattta atatgttata gctgaactctt ttgggtaact 420
ttaaattctt atcatagact ctgtacatat gttcaaatga gctgcttgcc tgatgtgtgt 480
atcatcggtg ggatgacaga acaaacatat ttatgatcat gaataatgtg ctttgtaa 538

```

&lt;210&gt; 312

&lt;211&gt; 176

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 312

```

ggaggagcag ctgagagata ggtgcagtga atgcgggtta gcctgctacc tctcctgtct 60
tcatagaacc attgccttag aattatttga tgacacgttt tttgttggtt aagctgttaag 120
gttttgttct ttgtgaacat ggttatattt aggggagggt ggaggagata gggaag 176

```

&lt;210&gt; 313

&lt;211&gt; 396

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens



```

<400> 313
ccagcacccc caggccctgg gggacctggg ttctcagact gccaaagaag ccttgccatc 60
tggcgctccc atggctcttg caacatctcc ccttcgtttt tgaggggggtc atgccggggg 120
agccacacag cctctactgg gttcgaggga gagtccaggaa gggccaagca cgacaaagca 180
gaaacatcgg atttggggaa cgcgtgtcaa tcccttgtgc cgcaggggctg ggcgggagag 240
actgtttctg tcttgtgtga actgtgttgc tgaagaacta cctcgttctt gtcttgatgt 300
gtcaccgggg ccaactgcctg ggggcgggga tgggggcagg tggaagcgg ctcgccattt 360
tataccaaag gtgctacatc tatgtgatgg gtgggg 396

```

```

<210> 314
<211> 311
<212> DNA
<213> Homo sapiens

```

```

<400> 314
cctcaacatc ctccagagagg actggaagcc agtccctacg ataaactcca taatttatgg 60
cctgcagtat ctcttcttgg agcccaaccc cgaggaccga ctgaacaagg agggccgcaga 120
ggtcctcgag aacaaccggc ggctgtttga gcagaacgtg cagcgctcca tgcgggggtg 180
ctacatcgcc tccacctact ttgagcgctg cctgaataag ggttgccgca taccaccccc 240
cgccacggcc acaagccctg gcatccctg caaatattta ttggggggcca tgggttagggg 300
tttggggggc g 311

```

```

<210> 315
<211> 336
<212> DNA
<213> Homo sapiens

```

```

<400> 315
tttagaacat ggttatcatc caagactact ctaccctgca acattgaact cccaagagca 60
aatccacatt cctcttgagt tctgcagctt ctgtgttaaa agggcagctg tcgtctatgc 120
cgtagaatca catgatctga ggaccattca tggaaagctgc taaatagcct agtctggggg 180
gtctttgata aagttttgca tggagcaaac aaacaggatt aaactaggtt tggttccctc 240
agccctctaa aagcataggg cttagcctgc aggccttcctt gggctttctc tgtgtgtgta 300
gttttgtaaa cactatagca tctgttaaga tccagt 336

```

```

<210> 316
<211> 436
<212> DNA
<213> Homo sapiens

```

```

<400> 316
aacatggtct gcgtgcctta agagagacgc ttctgcaga acaggacctg actacaaaga 60
atgtttccat tggaattggt ggtaaagact tggagttaac aatctatgat gatgatgat 120
tgtctccatt cctggaaggt cttgaagaaa gaccacagag aaaggcagag cctgctcaac 180
ctgctgatga aactgcagaa aaggctgatg aaccaatgga acattaagtg ataagccagt 240
ctatatatgt attatcaaat atgtaagaat acaggcacca cactactgat acaataatct 300
atactttgaa ccaaaagttg cagagtgggt gaatgctatg ttttaggaat cagtcagatc 360
gtgagttttt tcaagcaac ctcactgaaa cctatataat ggaatacatt tttctttgaa 420
agggctgtga taatca 436

```

```

<210> 317
<211> 196
<212> DNA
<213> Homo sapiens

```

```

<400> 317
tatcccttgt gaagatgata tactattttt gtttaagcgtg tctgtattta tgtgtgagga 60

```

```

gctgctggct tgcagtgcgc gtgcacgtgg agagctggtg cccggagatt ggaaggcctg 120
atgctccctc cctcgccctg gtccaggcaa gctggccgag ggtcctggct cctgaggggc 180
atctgccccct ccccca

```

<210> 318

<211> 381

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> 8, 9, 102, 122, 167, 182, 193, 235, 253, 265, 266, 290, 321, 378

<223> n = A,T,C or G

<400> 318

```

gacgcttngg ccgtaacgat gatcggagac atcctgctgt tcgggacgtt gctgatgaat 60
gccggggcgg tgcctgaact taagctgaaa aagaaggaca cncagggcct tggggaggga 120
tncaggggagc ccaacacagg tgacaacatc cgggaattct tgcctgacct cagatacttt 180
cnaatcttca tncocctgtg gaacatcttc atgatgttct gcatgattgt gctgntcggc 240
tcttgaatcc cancgatgaa accannaact cactttcccg ggatgccgan tctccattcc 300
tccattcctg atgactcaa naatgttttt gacacaaaaa ccgacaaact tcccagaaa 360
tccaagctcg tggtgggngg a

```

<210> 319

<211> 506

<212> DNA

<213> Homo sapiens

<400> 319

```

ctaagcttta cgaatggggg gacaacttat gataaaaaact agagctagtg aattagccta 60
tttgtaataa cctttgttat aattgatagg atacatcttg gacatggaat tgtaaagcca 120
cctctgagca gtgtatgtca ggacttggtc attaggttgg cagcagaggg gcagaaggaa 180
ttatcacagg agagatgtat gcagatgtgt ccataatatt ccatatttac atttttagat 240
ccattgatgt atgcattctt tggctgtact ataagaacac attaatcaaa tggaaataca 300
ctttgctaata attttaatgg tatagatctg ctaatgaatt ctcttaaaaa cactactgtat 360
tctgtgtcgt tegtgtttcat tttaaatgga gcattaaggg aatgcagcat ttaaatcaga 420
actctgccaa tgccttttato tagaggcgtg ttgccatttt tgccttatat gaaattttct 480
tcccaagaaa ggcaggatta catctt

```

<210> 320

<211> 351

<212> DNA

<213> Homo sapiens

<400> 320

```

ctgacctgca ggacgaaacc atgaagagcc tgatccttct tgccatcctg gccgccttag 60
cggtagtaac tttgtgttat gaatcacatg aaagcatgga atcttatgaa cttaatccct 120
tcattaacag gagaatgca aataccttca tatcccctca gcagagatgg agagctaag 180
tccaagagag gatccgagaa cgctctaagc ctgtccaaga gctcaatagg gaagcctgtg 240
atgactacag actttgcgaa cgctaagcca tggtttatgg atacaatgct gcctaataatc 300
gctacttcag gaagcgccga gggaccaaat gagactgagg gaagaaaaaa a 351

```

<210> 321

<211> 421

<212> DNA

<213> Homo sapiens

```

<400> 321
ctcggaggcg ttcagctgct tcaagatgaa gctgaacatc tccctcccag ccactggctg 60
ccagaaactc attgaagtgg acgatgaacg caaacttcgt actttctatg agaagcgatg 120
ggccacagaa gttgctgctg acgctctggg tgaagaatgg aaggggtatg tggccgaat 180
cagtggtggg aacgacaaac aaggtttccc catgaagcag ggtgtcttga cccatggccg 240
tgtccgctgt ctactgagta aggggcattc ctgttacaga ccaaggagaa ctggagaaag 300
aaagagaaaa tcagttcgtg gttgcattgt ggaagcaaat ctgagcgttc tcaacttggt 360
tattgtaaaa aaaggagaga aggatattcc tggactgact gatactacag tgccctcgcc 420
c 421

```

<210> 322

<211> 521

<212> DNA

<213> Homo sapiens

```

<400> 322
agcagctctc ctgccacagc tctctacccc ctgaaaatgt tgcgctgctc caagtttgtc 60
tccactccct ccttggtcaa gagcaacctc cagctgctga gccgtccgct atctgcagtg 120
gtgctgaaac gaccggagat actgacagat gagagcctca gcagcttgcc agtctcatgt 180
ccccttacct caactgtctc tagccgcagc ttccaaacca gcgcatttc aagggacatc 240
gacacagcag ccaagttcat tggagctggg gctgccacag ttgggggtggc tgggttctggg 300
gctgggattg gaactgtggt tgggagccct atcatttggt atgccagaaa cccctctctg 360
aagcaacagc tcttctccta cgccattctg gcctttgccc tctcggagcc catggggctc 420
ttttgtctga tggtagcctt tctcatctct tttgccatgt gaaggagccg tctccacctc 480
ccatagttct cccgcgtctg gttggccccg tgtgttctct t 521

```

<210> 323

<211> 435

<212> DNA

<213> Homo sapiens

```

<400> 323
ccgaggtcgc acgcgtgaga cttctccgcc gcagacgcgg ccgcgatgag ctacgtcgcc 60
tctacactgc tggtgcctct agggggcaac tctcccccga gccccaagga catcaagaag 120
atcttgaca gcgtgggtat cgaggcgagc gacgaccggc tcaacaaggt tatcagtga 180
ctgaatggaa aaaacattga agacgtcatt gccagggtta ttggcaagct tgccagtgta 240
cctgctggtg gggctgtagc cgtctctgct gccccagcgt ctgcagcccc tgctgctggt 300
ctgccccctg ctgcagcaga ggagaagaaa gatgagaaga aggaggagtc tgaagagtc 360
gatgatgaca tgggatttgg cctttttgat taaattctgt ctcccctgca aataaagcct 420
ttttacacat ctcaa 435

```

<210> 324

<211> 521

<212> DNA

<213> Homo sapiens

```

<400> 324
aggagatoga ctttcggtgc ccgcaagacc agggctggaa cgccgagatc acgtgcaga 60
tggtgcagta caagaatcgt caggccatcc tggcggtcaa atccacgcgg cagaagcagc 120
agcaccttgt ccagcagcag cccccctcgc agccgcagcc gcagccgcag ctccagcccc 180
aaccocagcc tcagcctcag ccgcaacccc agcccccaat acaaccccag cctcagcccc 240
aaccocagcc tcagccccag cagctccacc cgtatccgca tccacatcca catccacact 300
ctcatctcta ctgcagccca caccctcacc cgcaaccgca tccgacacaa ataccgcacc 360
cacaccacaa gccgcactcg cagccgcagc ggccaccgct tctccgcagc acctccaaact 420
ctgcctgaaa ggggcagctc ccgggcaaga caaggttttg aggaattgag gaagtgggac 480
gagcacattt ctattgtctt cacttggtac aaaagcaaaa c 521

```

<210> 325

<211> 451  
 <212> DNA  
 <213> Homo sapiens

```
<400> 325
attttcattt ccattaacct ggaagctttc atgaatatcc tttcttttta aaacatttta 60
acattatttta aacagaaaaa gatgggctct tctcggttag ttgttacatg atagcagaga 120
tatttttact tagattactt tgggaatgag agattgttgt cttgaaactc ggcactgtac 180
agtgaatgtg tctgtagtgt tggttagttg cattaaagcat gtataacatt caagtattgtc 240
atccaaataa gaggcataata cattgaattg tttttaatcc tctgacaagt tgactcttgc 300
acccccacc ccaccaaga cattttaata gtaaatagag agagagagaa gagttaatga 360
acatgaggta gtgttccact ggcaggatga cttttcaata gctcaaatca atttcagtcg 420
ctttatcact tgaattatta acttaatttg a 451
```

<210> 326  
 <211> 421  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> 296  
 <223> n = A,T,C or G

```
<400> 326
cgcggtcgta agggctgagg atttttggtc cgcacgctcc tgctcctgac tcaccgctgt 60
tcgctctcgc cgaggaacaa gtcggtcagg aagcccgccg gcaacagcca tggcttttta 120
ggataccgga aaaacacccg tggagccgga ggtggcaatt caccgaattc gaatcacctc 180
aacaagccgc aacgtaaaaa ccttgaaaaa ggtgtgtgct gacttgataa gaggcgcaaa 240
agaaaaaagt ctcaaatgta aaggaccagt tcgaatgctt accaagactt tgagantcac 300
tacaagaaaa actcctctgt gtgaagggtc taagacgtgg gatcgtttcc agatgagaat 360
tcacaagcga ctcattgact tgcacagctc tctgagatt gttaagcaga ttacttccat 421
c
```

<210> 327  
 <211> 456  
 <212> DNA  
 <213> Homo sapiens

```
<400> 327
atcttgacga ggctgcggtg tctgctgcta ttctccgagc ttgcgaatgc cgcttaagga 60
cgacaagaag aagaaggagc ctggaagatc ggccaagaaa gacaaagacc cagtgaacaa 120
atccgggggc aaggccaataa agaagaagtg gtccaaaggc aaagtctcgg acaagctcaa 180
taacttagtc ttgtttgaca aagctaccta tgataaactc tgtaaggaa gttcccaacta 240
taaacttata accccagctg tggctcttga gagactgaag attcgaggtc cctctggccag 300
ggcagccctt caggagctcc ttagttaaag acttatcaaa ctggtttcaa agcagcagac 360
tcaagtaatt tacaccagaa ataccaaggg tggagatgct ccagctgctg gtgaagatgc 420
atgaataggt ccaaccagt gtacattttg aaaaaa
```

<210> 328  
 <211> 471  
 <212> DNA  
 <213> Homo sapiens

```
<400> 328
gtggaagtga catcgtcttt aaacctgcgc tggcaatccc tgacgcacgc ccgtgatgcc 60
cagggaagac agggcgacct ggaagtccaa ctacttccct aagatcatcc aactattgga 120
tgattatccg aaatgtttca ttgtgggagc agacaatgtg ggtccaagc agatgcagca 180
```

```

gatcgcgatg tcccttcgcg ggaaggctgt ggtgctgatg ggcaagaaca ccatgatgcg 240
caaggccatc cgagggcacc tggaaaaaca cccagctctg gagaaactgc tgccctcatat 300
cggggggaat gtgggctttg tgttcaccaa actgagatca gggacatgtt 360
gctggccaat aaggtgccag ctctgtcccg tgcgtgtgac attgcccacat gtgaagtcaac 420
tgtgccagcg cagaacactg gtctcgggcc cgagaagacc tcctttttcc a 471

```

&lt;210&gt; 329

&lt;211&gt; 278

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; 154, 204

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 329

```

gtttaaactt aagcttggtg ccgagctcgg atccactagt ccagtgtggt ggaattctag 60
aaattgagat gccccccacc gccagcaaat gttccttttt gtccaagatc tatttttatt 120
ccttgatatt tttctttttt tttttttttt ttgnggatgg ggaacttgtg atttttctaa 180
aggtgctatt taacatggga gganagcgtg tgcggctcca gccacgcccg ctgctcactt 240
tccaccctct ctcaacctgc ctctggcttc tcaggcct 278

```

&lt;210&gt; 330

&lt;211&gt; 338

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 330

```

ctcaggcttc aacatgaat acgcgcgagg ccccttcgcc ctattcttca tagccgaata 60
cacaacattt attataataa acaccctcac cactacaatc ttcctaggaa caacatatga 120
cgcactctcc cctgaactct acacaacata ttttgtcacc aagaccctac tcttaacctc 180
ccttgattct tgaattcgaa cagcatatccc ccgattccgc tacgaccaac tcataacctc 240
cctatgaaaa aactctctac cactcaccct agcattactt atatgatgat tctccatacc 300
cattacaatc tccagcattc cccctcaaac ctaaaaaa 338

```

&lt;210&gt; 331

&lt;211&gt; 2820

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 331

```

tggcaaaatc ctggagccag aagaaaggac agcagcattg atcaatctta cagctaacat 60
gttgtaacct gaaaacaagt cccagactca atttagtgag ccacagtaca cgaacctggg 120
gctcctgaac agcatggacc agcagattcg gaacggctcc tcgtccacca gtccctataa 180
cacagaccac gcgcagaaca gcgtcacggc gccctcgccc tacgcacagc cagcccccac 240
cttcgatgct ctctctccat caccgcccat cccctccaac accgactacc caggcccaga 300
cagttccgac gtgtccttcc agcagtcgag caccggccaa tcggccacct ggacgtatcc 360
cactgaactc aagaaactct actgcacaat tgcaaaagca tgcccacatc agatcaagg 420
gatgaaccga cctcctcagg gagctgttat ccgcgccatg cctgtctaca aaaaagctga 480
gcacgtcacg gaggtggtga agcgggtgcc caaccatgag ctgagccgtg agttcaacga 540
gggacagatt gccctccta gtcatattgat tcgagtagag ggaacacagc atgcccagta 600
tgtagaagat cccatcacag gaagacagag tgtgctggta ccttatgagc caccgccagt 660
tggaactgaa ttccagacag tcttgtacaa tttcatgtgt aacagcagtt gtgttggagg 720
gatgaaccgc cgtccaattt taatcattgt tactctggaa accagagatg ggcaagctct 780
gggcccagcg tgctttgagg cccggatctg tgcttgcaca ggaagagaca ggaaggccga 840
tgaagatagc atcagaagcg agcaagtttc ggacagtaca aagaacgggt atggtacgaa 900
gcgcccggtt cgtcagaaca cacatgggat ccagatgaca tccatcaaga aacgaagatc 960

```

```

cccagatgat gaactgttat acttaccagt gaggggccc gagacttat aaatgctgtt 1020
gaagatcaaa gagtccctgg aactcatgca gtaccttctt cagcacacaa ttgaaacgta 1080
caggcaacag caacagcagc agcaccagca cttacttcag aacacagcct caatacagtc 1140
tccatcttca tatggtaaca gctccccacc tctgaacaaa atgaacagca tgaacaagct 1200
gcctctgtg agccagctta tcaaccctca gcagcgcaac gccctacct ctacacacct 1260
tctctgatgg atgggagcca acattcccat gatgggcaac ccacatgcaa ttgctggaga 1320
catgaatgga ctacagcccca cccaggcact cctcctccca ctctccatgc catccacctc 1380
ccactgcaca cccccacct cgtatccac agattgcagc attgtcagtt tcttagcgag 1440
gttgggctgt tcatcatgtc tggactattt cagcaccagc gggctgacca ccatctatca 1500
gattgagcat tactccatgg atgatctggc aagtctgaaa atccctgagc aatttcgaca 1560
tgcatctgg aagggcatcc tggaccacgc gcagctccac gaattctcct ccccttgcta 1620
tctcctcggc accccaagca gtgcctctac agtcagttgt ggctccagtg agaccggggg 1680
tgagcgtgtt attgatctgc tgcgattcac cctccggcag accatctctt tcccaccccg 1740
agatgagtg aatgaactca actttgacat ggatgctcgc cgcaataagc aacacgcgat 1800
caaaagaggg ggggagtga cctcaccatg tgagctcttc ctatccctct cctaaactgc 1860
agccccctaa aagcactcct gcttaactct caaagccttc tccctagtct cctcccttcc 1920
tcttgtctga tttcttaggg gaaggagaag taagaggcta cctcttaacct aacactctag 1980
ctggcgtcta attctgattc tggctttaag ccttcaaaac tatagcttgc agaactgtag 2040
ctgccatggc taggtagaag tgagcaaaaa agagttgggt gtctccttaa gctgcagaga 2100
tttctcattg acttttataa agcatgttca cccttatagt ctaagactat atataataa 2160
tataaatatg acagtataga tttttgggtg gggggcattg agtattgttt aaactgtatt 2220
ttaaataaaa gaaaattgag ttgcacttat tgaccatttt ttaatttact tgttttggtat 2280
ggctgtgcta tactccttgc cttaaagggt atcatgtatg gtgatagcta ctagagctt 2340
aatgctacat gtgagtgcga tgatgtacag attcttctag tcttttggtat tctaataaca 2400
tgccacatca aacctttgag tagatccatt tccattgctt attatgtagg taagactgta 2460
gatagtctat cttttctcag ttgttggtata ttttatatta ctgacatttc tctagtgtat 2520
gatggtctac gttggggtga ttaatccag ttataagaag aagttcatgt ccaaacggtc 2580
ctctttagtt ttgggttggg aatgaggaaa attcttataa agggcatagc agccagttca 2640
aaaaacccgc acgtcatgta tttgagcata tcagtaaccc ccttaaat ttaataccaga 2700
taccttatct tacaatgttg attgggaaaa catttgtctg ccattacaga ggtattaaaa 2760
ctaaatttca ctactagatt gactaactca aatacacatt tgctactgtt gtaagaattc 2820

```

&lt;210&gt; 332

&lt;211&gt; 2270

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 332

```

tcgttgatga caaagacagt tgaaggaaat gaattttgaa acttcacggg gtgccaccct 60
acagtactgc cctgacctt acatccagcg tttcgtagaa acccagctca tttctcttgg 120
aaagaaagttt attaccgcat caccatgtcc cagagcacac agacaaatga attcctcagt 180
ccagaggtttt tccagcatat ctgggatttt ctggaacagc ctatatgttc agttcagccc 240
attgacttga actttgtgga tgaaccatca gaagatggtg cgacaaacaa gattgagatt 300
agcatggact gtatccgatc gcaggactcg gacctgagtg accccatgtc gccacagtac 360
acgaacctgg ggcctcgtga cagcatggac cagcagattc agaacggctg gctctcacc 420
agtcctcata acacagacca cgcgcagaa acgctcacgg cgccctcgcc ctacgcacag 480
cccagctcca cettcgatgc tctctctcca tcaccgcga tccctccaa cacogactac 540
ccaggcccgcc acagtttcca cgtgtcctc cagcagtcga gcaccgcaa gtgcggcacc 600
tggaagtatt ccactgaact gaagaaactc ttggtccaaa ttgcaaaagac atgcccctac 660
cagatcaagg tgatgacccc acctcctcag gtagctgtta tcgcgcgctc cctgtctcac 720
aaaaaagctg agcacgtcac ggaggtgggt agcgggtgcc ccaaccatga gctgagccgt 780
gaattcaacg agggacagat tgcccctcct aatcatttga ttcagtaga ggggaactac 840
catgccacgt atgtagaaga tcccatcaca ggaagacaga gtgtgctggt accttatgag 900
ccaccccgag ttggcactga attcacgaca ctcttgata atttcatgtg taacacagat 960
tgtgttgagg ggatgaacgg cgtccaatt ttaactcatt ttaactctga aaccagagat 1020
gggcaagtgc tgggcccagc ctgcttttag gcccgatct gtgcttgccc aggaagagac 1080
aggaaggcgg atgaagatag catcagaaag cagcaagttt cggacagtac aaagaacggt 1140

```

gatggtacga	agcgcccggt	tgcgcagaac	acacatggta	tccagatgac	atccatcaag	1200
aaacgaagat	ccccagatga	tgaactgtta	tacttaccag	tgaggggccg	tgagacttat	1260
gaatgtcgtg	tgaagatcaa	agagtcctgt	gaactccttc	agtcacttcc	tcagcacaca	1320
ttatgaacgt	acaggcaaca	gcaacagcag	cagcaccagc	acttacttca	gaacacagacc	1380
tcaatcacgt	ctccatcttc	atatggtaac	agctcccac	ctctgaacaa	aatgaacagc	1440
atgaacaacg	tgcctctctg	gagccagctt	atcaaccctc	agcagcgcaa	cgccctcaact	1500
cctacaacca	ttcctgatgg	catggggagcc	aacattccca	tgatgggac	ccacatgcca	1560
atggctggag	acatgaatgg	actcagcccc	accggagcac	tcctctcccc	gctctccatg	1620
ccatccacgt	cccactgcac	acccccacct	ccgtatccaa	cagatgtcag	cattgtcggt	1680
ttcttagcga	ggttggcggt	ttcatcatgt	ctggactatt	tcacgaccca	gggctgacc	1740
accatctatc	agattgagca	ttaactccatg	gatgatctgg	caagtctgaa	aatccctgag	1800
caatttcgac	atgcgatctg	gaagggcatc	ctggaccacc	ggcagctcca	cgaattctcc	1860
tcccctcttc	atctcctgag	gaccccaagc	agtgcctcta	cagtcatgtg	gggctccagt	1920
gagaccggg	gtgagcgtgt	tattgatgct	gtgcgatcca	ccctccgcca	gaccatctct	1980
ttccccccc	gagatgagtg	gaatgacttc	aactttgaca	tggatgctcg	ccgcaataag	2040
caacagcgca	tcaaaagagga	gggggagtgga	gcctcaccat	gtgagctctt	ccatccctc	2100
tcttaaacctg	cagccccccta	aaagcactcc	tgcttaatct	tcaaaagcctt	ctccctgagt	2160
cctccccttc	ctcttgcttg	atttctttag	ggaaaggagaa	gtcgaaggct	acctcttacc	2220
taacatctga	cctggcatct	aattctgatt	ctggcttttaa	gccttcaaaa		2270

&lt;210&gt; 333

&lt;211&gt; 2816

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 333

tcggtgatata	caaagacagt	tgaaggaaat	gaattttgaa	acttcacggt	gtgccaccct	60
acagtatctgc	cctgaccctt	acatccagcg	ttctgtagaa	acccagctca	ttctctttgg	120
aaagaaagtt	attaccgatc	caccatgtcc	cagagcacac	agacaatatga	attcctcagt	180
ccagaggttt	tccagcatat	ctgggatttt	ctggaaacagc	ctatatgttc	agttcagccc	240
attgacttga	actttgtgga	tgaaccatca	gaagatggtg	cgacaaacaa	gattgagatt	300
agcatggact	gtatcccgat	gcaggactcg	gacctgagtg	acccccatgtg	gccacagatc	360
acgaaacctgg	ggctcctgaa	cagcatggac	cagcagattc	acacgcgctc	ctcgtccacc	420
agtcctcata	acacagacca	cgcgcagaac	agcgtcacgg	cgccctcgcc	ctacgcacag	480
cccagctcca	ccttcgatag	tctctctcca	tcacccgcca	tcctctccaa	caccgactac	540
ccaggccccc	acagtttoga	cgtgtctctc	cagcagtcga	gcaccgcaa	gtcgccacc	600
tggacgtatt	ccaactgaact	gaagaaactc	tactgccaaa	ttgcaaaagc	atgccccatc	660
cagatcaagg	tgatgacccc	acctcctcag	ggagctgtta	tcgcgcgcac	gcctgtctac	720
aaaaaagctg	agcacgtcac	ggaggttggtg	aagcgggtgcc	ccaacataga	gctgagccgt	780
gaattccaag	agggacagat	tgcccctcct	agtcatttga	ttcgaagtaga	aaagaacggt	840
catgccacgt	atgtagaaga	tcccatcaca	ggaagacaga	gtgtgctggt	accttatgag	900
ccaccccag	tgggactgga	attcacgaca	gtcttgata	atttcatgtg	taacagcagt	960
tgtgttggag	ggatgaaccg	cogtccaat	ttaatcatgt	ttactctgga	aaccagagat	1020
gggcaagctg	tggggcgcag	ctgctttgag	gcccgagatc	gtgcttgccc	aggaagacag	1080
aggaaggcgg	atgaagatag	catcagaaga	cagcaagttt	cggacagtat	aaagaacggt	1140
gatggtacga	agcgcccggt	tgcgcagaac	acacatggta	tccagatgac	atccatcaag	1200
aaacgaagat	ccccagatga	tgaactgtta	tacttaccag	tgaggggccg	tgagacttat	1260
gaatgtcgtg	tgaagatcaa	agagtcctgt	gaactcatgc	agtaaccttc	tcagcacaca	1320
ttatgaacgt	acaggcaaca	gcaacagcag	cagcacagc	acttacttca	gaacatcttc	1380
cttttcacgt	gcttcagaaa	tgagctttgtg	gagccccgga	gagaaactcc	aaacacactc	1440
gacgtcttct	ttagacattc	caagccccc	aaccgatcag	tgtaccata	gagccctatc	1500
tctatatttt	aagttgtgtg	gttgtatttc	catgtgtata	tgtgagtgtg	tgtgtgtgta	1560
tctgtgtctg	tgtgtatctc	gccctcataa	acaggacttg	aagacacttt	gtcgcagaca	1620
cccacactgt	caaaggcaca	aagccactag	tgagagaatc	ttttgaaggg	actcaaacct	1680
ttacaagaaa	ggatgttttc	tgcaagattt	gtactcttag	accggccatt	ggtgggtgag	1740
gaaccactgt	tttgtctctg	gagctttctg	ttgtttctcg	ggagggagg	gtcaggtggg	1800
gaaaggggca	ttaagatgtt	tattggaacc	ctttctgtct	ttcttctggt	gtttttctaa	1860
aattcacag	gaagcttttg	agcaggtctc	aaactaaga	tgtcttttta	agaaaaggag	1920

aaaaaagtgtg	ttattgtctg	tgcataagta	agttgttaggt	gactgagaga	ctcagtcaga	1980
cccttttaaat	gctgggtcatg	taataatatt	gcaagtagta	agaaaacgaag	gtgtcaagtg	2040
tactgtctggg	cagcgaggtg	atcattacca	aaagtaatca	acctttgtggg	tggagagttc	2100
tttgtgagaa	cttgcatatt	ttgtgtcctc	ccctcatgtg	taggtagaac	atttcttaat	2160
gctgtgtacc	tgctctgcc	actgtatgtt	ggcatcgttt	atgctaaaag	ttttcttgta	2220
catgaaaccc	tggaagacct	actacaaaaa	aactgttgtt	tggcccccac	agcaggtgaa	2280
ctcatttttgt	gcttttaata	gaagacaaaa	tcacccccag	taatatgtcc	cttactagtg	2340
tgttttaccat	tattcaaaagc	tcaaaataga	atttgaagcc	acctcacaaa	actctgtagt	2400
aaattgtctta	atttagagctt	ctatccctca	agcctaccta	ccataaaaacc	agccataatta	2460
ctgatacgtt	toagtgctatt	tagccaggag	acttacctgt	tgagtaagtg	agccaagaag	2520
agacgtgtta	aaatcagcac	tctctggactg	gaaattaaag	attgaaaggg	tagactactt	2580
ttctttttttt	tactcaaaaag	tttagagaat	ctctgtttct	ttccatttta	aaaacataatt	2640
taaagataat	agcataaaaga	ctttaaaaat	gttctctccc	tcactcttcc	cacacccagtg	2700
caccagcaact	gtattttctg	tcaccaagac	aatgatttct	tgttattgag	gctgtgtcct	2760
tttgtgatgt	gtgattttta	ttttcaataa	acttttgcat	cttggtttta	aagaaa	2816

&lt;210&gt; 334

&lt;211&gt; 2082

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 334

agatgctaca	gogactgcac	acccaggctg	tatgatcacg	ccattgtctc	cogggctgca	60
aacctgtcca	gcattgtgatg	tggtgggata	ctgaattgaa	taccgaatac	tgtaggccaat	120
tgtaaacacag	tggttaagctc	ttgtgtatct	aaacatagct	aaacacaaa	aggtatagta	180
agaatatgtgt	altataatct	tatggaacta	tcattgtata	tgtggtttgt	aacccagaat	240
gtagttatata	agcacaggagc	tggtgcttatg	atgtgccaaag	cacagctctc	agtcataact	300
cccttaaatct	tcataataac	cctaggaggt	aaactcttaa	gtagattcat	attgtaagggt	360
ctctcgggggt	gggggggttgg	caaaatcctg	gagccagaag	aaaggacagc	agcatgtatc	420
aatcttcacg	ctaacatgtt	tgacctggaa	aaacatgcc	agactcaatt	tgtgtagcca	480
cagtacacga	acctgggggt	cctgaacagc	atggaccagc	agattcagaa	cggctcctcg	540
tcacacagct	ccataaacac	agaccacgag	cagacacagc	tcacggcgcc	ctcgcctcac	600
gcacagaccca	gctccacott	cgatgctctc	tctccatcac	ccgccatccc	ctccaacacc	660
gactacctcga	gccgcacagc	tttcgacgtg	tccttcacag	agtcgagcac	cgcccaagtg	720
gccaccccgga	cgtattccac	tgaactgaag	aaactctact	gccaaattgc	aaagacatgc	780
cccatccaga	tcaaggtgat	gacccacact	cctcagggag	ctgttatccg	cgccatgcct	840
gtctacaaaa	aagctgagca	cgctcacggag	gtggtagaag	gggtgcccaa	ccatgagctg	900
agccgtgtaat	tcaacgaggg	acagattgcc	ccctcagtc	atttgattcg	agtagagggg	960
aaacgccaat	cccagtatgt	agaagatccc	atcacaggaa	gacagagtgt	gctggtacct	1020
tatgagccac	cccaggttgg	caactgaatt	acgacagctc	gtacaatttc	catgtgtaac	1080
agcagtttgt	ttggagggat	gaaccgcgct	ccaattttta	tcattgtttac	cttggaaac	1140
agagatgggc	aagtcctggg	ccgacgctcg	tttgaggccc	ggatctgtgc	ttgcccagga	1200
agagacagga	agggcgatga	agatagcatc	agaaagcagc	aagtttcgga	cagtacaaaag	1260
aacggtgatgc	gtacgaagcg	cccgctctcg	cagaacacac	atggtatcca	gatgacatct	1320
actaagaatac	gaagatcccc	agatgatgaa	ctgttatact	taccagtgag	ggggcggtgag	1380
accttatgaa	tgctgttgaa	gatcaaaag	tccttggaa	tcatgcagta	cccttctcag	1440
cacacaattg	aaacgtacag	gcaacagcaa	cagcagcagc	accagcaact	acttcagaaa	1500
cagtgagtgt	atcaactgtgt	cattttagga	ggcatgagtg	acggtgcact	tatttggatc	1560
agcaataagg	tgatttgatga	gcaatgtgga	acataatggg	agatagcaga	ttgtcataga	1620
ttcagatgac	ctggtatggc	aaccctcttt	cagtttgcaac	cttttttaag	tgctctatta	1680
taacctcttc	ctcagaattc	cacttatgtt	ctgaaattaa	atacaaacca	tttctggtga	1740
attacaaga	aactcacact	aacagttctc	ttctctatat	gcttggtcca	ataacactaa	1800
cagtagatgac	acactctatt	tggttagtgat	gtgtatatatt	gaaaacacga	tactctttct	1860
catcccaatg	gattgtctta	taaatctcct	gggatgcaca	ctatccactt	ttgggaataa	1920
ctgtgtagac	cagggtatgc	aaatagggtt	tactataata	taaggtgact	gtttggaatg	1980
ctgtaattgag	aagaattctg	agacctatgt	catgataatt	ggggaaatat	gtgggtgcag	2040
aaggataaag	tagcatcatg	ttgccttatt	ttagcatctc	tg		2082



&lt;210&gt; 335

&lt;211&gt; 4849

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 335

```

cggtgatatc aaagacagtt gaaggaaatg aattttgaaa cttcacgggtg tgccacccta 60
cagtagatgcc ctgaccctta tctctagaaa cccacagctca ttctctctgg 120
aaagaaagttt attaccgatc caaccatgtcc cagagcacac agacaaatga attcctcagt 180
ccagagggttt tcacgcatat ctggggtttt ctggaacacg ctatatgttc agttcagccc 240
attgacttga actttgtgga tgaacctaca gaagatgggt cgacaaacaa gatttgagatt 300
agcatggact gtatccgcat gcaggactcg gacctgagt accccatgtg gccacagtc 360
acgaacctgg ggctcctgaa cagcatggac cagcagatto agaacggctc ctgctccacc 420
agtcctcata acacagacca cgcgcagaac agcgtcacgg cgccctcgcc ctacgcacag 480
ccagcttcca ccttcgatgc tctctctcca tcacccgcga tccctcccaa ccccgactac 540
ccaggccggc acagttttga cgtgtccttc cagcagtcga gccccgccaa ctgggccacc 600
tggaagtatt ccaactgaatc gaagaaactc tactgccaaa ttgcaaaagc attccccatc 660
cagatcaagg tgatgacccc acctcctcag ggagctgtta tcgccgccac gctgtcttac 720
aaaaaagctg agcacgtcac ggaggtgggt aagcgggtgc ccaaccatga gctgagccgt 780
gaattcaacg agggacagtt tgccctcctc agtcatttga ttcagataga ggggaacagc 840
catgcccagt atgtagaaga tcccatcaca ggaagacaga gtgtgtctgt acctttagag 900
ccaccccagg ttggcactga attcacgaca gtctgttaca atttcagtgt taacagcagt 960
tgtgtttgag ggaatgacccg ccgtccaatt ttaactcattg ttactctgga aacacagagt 1020
gggcaagtcc tgggcgcagc ctgcttttag gcccgatct gtgcttgccc aggaagagac 1080
aggaaggcgg atgaagatag catcagaagg cagcaagttt cggacagtac aaagaacggt 1140
gatggtaaga agcgcocggt tcgtcagaac acacatggta tccagatgac atccatcaag 1200
aaacgaagat cccacagatga tgaactgtta tacttaccag tgaggggccc tgagacttat 1260
gaattgctgt tgaagtcaca agagtccctg gaactcatgc agtacccttc tcagcacaca 1320
attgaaacgt acaggcaaca gccacagcag cagcacccagc acttactctga gaaacagacc 1380
tcaatcacgt tccatctctc atatggtaaac agctccccc ctctgaacaa aatgaacagc 1440
atgaacaagg tgccttctgt gagccagctt atcaacccctc agcagcgcaa cgccctcact 1500
cctacaacca ttctgatgg catgggagcc aacattccca tgatgggcac acatcagcca 1560
atgggtggag acatgaatgg actcagcccc acccagggac tcocctcccc cctctccatg 1620
ccatccacct cccagtgcac acccccacct ccgtatccca cagattgcag acctgtcagt 1680
ttcttagcga ggttgggctg ttcatcatgt ctggactatt tcacgaccca ggggctgacc 1740
accatctatc agattgagca ttactccatg gatgatctgg caagctgtaa aatccctgag 1800
caattctcag atcgcatctg gaagggcata ctggacaccc ggcagctctca cgaattctcc 1860
tcccctctgc atctcctgcg gaocccaaagc agtgctctta cctctcgcca gaccatctct 1920
gagaccccgg gtgagcgtgt tattgatgct gtgcgattca cctctcgcca gaccatctct 1980
ttccaccccc gagatgagtg gaatgacctc aactttgaca ttgatctgga ccgcaataag 2040
caacagcgca tcaaaagaga gggggagtgga gccctaccat gtgagctctt cctatccctc 2100
tctaatcagt cagcycccta aaagcactcc tgcattatct tcaaaagcctt ctccctagct 2160
cctcccccctg ctcttctctg atttcttagg ggaaggagaa gtaagaggt accctttacc 2220
taacatctga cctggcatgc aattctgatt ctggctttaa gccttcaaaa ctatagcttg 2280
cagaactgta cgtgccatgg ctaggtagaa gacttttata aagcatgttc acccttaatt gctgtcctta 2340
agctgcagag atttctcatt tacagtatag atttttgggt gggggcattg tctaagacta 2400
tatataataa tgtataataa tgcataatga atttttgggt gggggcattg gatttatgtt 2460
taaaatgtaa ttttaaatgaa agaaaattga gttgcactta ttgacatttt tttatttttt 2520
ttgttttggg tggcttctct atactccttc ccttaagggg tatcatgtat ggtgataggt 2580
atctagagct taatgctaca tgtgagtgc gatgatgac agattctctt agtcttttgg 2640
attctaaata catgcacact caaacctttg agtagatcca ttcccatatg ttattatgta 2700
ggtaagactg tagatatgta ttcttttctc agtgttggtta tattttatgc tactgaocat 2760
tctctagtg atgatgtgtc acgttggggt gatttaatcc agttataaga agaagttcat 2820
gtccaaacgt cctctttagt ttttggttgg gaatgggaa aattcttaaa agggccatag 2880
cagccagctg aaaaacaccc gacgtcatgt atttgagcat atcagtaacc cctttaaatt 2940
taataccaga taccttatct tacaatatgt attgggaaaa catttgctgc cattacagag 3000
gtattaaaaa taaatttccac tactagattg actaaactca atacactgtt gctactgttt 3060
taagaattct gatgatttgg attgggatga atgcacatca tctagttcta acagtgaagt 3120

```

tttactgtct	attaatatct	agggtaaata	ggaatcattc	agaaatgttg	agtctgtact	3180
aaacagtaag	atatctcaat	gaaccataaa	ttcaactttg	taaaaatctt	ttgaagcata	3240
gataatattg	tttggtaaat	gtttcttttg	tttggtaaat	gtttctttta	aagaccctcc	3300
tattctataa	aactctgcat	gttagagctt	gtttacottt	ctctctctaa	ggttttacaat	3360
aggagtgtgt	atttgaaaaa	tataaaatta	tgagatttgt	tttctgtgtg	catataattgc	3420
atcaactgtat	cattttctttt	tttaaccggg	aagagtttca	gtttgtgtga	aagtaactgt	3480
gagaaccocag	tttcccgctc	atctccctta	gggactacc	atagacatga	aaggtcccca	3540
cagagcaaga	gataagctct	tcatggctgc	tggttgctta	accacttaaa	cgaagagttc	3600
ctttgaaact	ttgggaaaac	atgttaatga	caatatcca	gatctttcag	aaataataca	3660
cattttttttg	catgcatgca	aatgagctct	gaaactctcc	catgcatctc	ggtaaggggc	3720
tgctattgca	cataagcttc	catttttaatt	ttaaagtgc	aaagggccag	ctgtgctcta	3780
aaaggtaaatg	tgtgatttgc	ctctgaaaag	tgtgtatata	ttttgtgtga	aattgcatac	3840
tttgtattttt	gattattttt	tttttcttct	ttggatagtg	ggatttccag	aaccacactt	3900
gaaaacttttt	tttatcgttt	ttgtattttc	atgaaaatac	cattttagtaa	gaataccaca	3960
tcaaaataaga	aataatgcta	caatttttaag	aggggaggga	agggaaagtt	ttttttttat	4020
atattttttat	aatttttgtat	gttaaagaga	atgagctctt	gatttcaaat	ttttgtgtga	4080
cttaaatggt	aataagcaat	gtaaaactct	gcaacaagca	tgagctttg	caaacccatt	4140
aaggggaaga	atgaagcgtg	ttccttggtc	ctagtaagaa	gacaaactgc	ttcccttact	4200
ttgctgaggg	tttgaataaa	ctcaggactt	ccgagctatg	tcagtactat	tcaggttaaca	4260
ctagggcctt	ggaaattcct	gtactgtgtc	tcactggatc	ggcactagcc	aagacggagg	4320
acccttaactg	gcttaccctc	tcactggcagc	ctactctcct	tgagtgtatg	atgagccagg	4380
gtaaggggtga	aaaggatagt	aagcatagaa	accactagaa	agtggtgtta	atggagttct	4440
tgtgtgctcga	gtctaatgca	gttagctgaa	gaattgaaaa	gttttgtgtt	ggagagcgtt	4500
ataaacacgaa	atggaagaca	gagttttcat	taaatccttt	tacotttttt	ttttcttggt	4560
aatccctcata	aataacaaat	tgtgggatat	tgaatgttaa	agggatattt	tttttctatt	4620
atttttataat	ttgtacaaaa	ttaaagcaat	gttaaaagtt	ttatatgctt	tattaatgtt	4680
ttcaaaaggt	attatacatg	tgatacattt	tttaagcttc	agttgtgttt	ctctgtgtac	4740
tttctgtatg	gggcttttgg	ggagccagaa	gccaatctac	aatctctttt	tgtttgccag	4800
gacatgcaat	aaaatttaaa	aaataataaa	aaactaatta	agaaataaaa		4849

&lt;210&gt; 336

&lt;211&gt; 1386

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 336

atgttgtaacc	tggaatacaa	tgcccagact	caatttagtg	agccacagta	caogaacctg	60
gggctccctga	acagcatgga	ccagcagatt	cagaacggct	cctcgtccac	cagtccctat	120
aaacacagacc	acgcgcagaa	cagcgtcacg	gcgcctctgc	ctacgcgaca	gccagcgtcc	180
acctctcgatg	ctctctctcc	atcacccgcg	atccctctca	acacgcgacta	cccgagcccg	240
cacagtttctg	acgtgtcctt	ccagcagtcg	agcaccgcga	agtcggccac	ctggacgtat	300
tcacatgaac	tgaagaaact	ctactgcaca	attgcaaaag	cagcccccac	ccagatcaag	360
gtgatgacc	caactctcca	gggagctgtt	atccgcgcga	tgctgtctca	caaaaaagct	420
gagcagctga	cggaggttgt	gaagcgtgtc	cccaaccatg	agctgagccg	tgaattcaac	480
gagggacaga	ttgcccctcc	tagtcaattg	attcgagtag	aggggaacag	ccatggccag	540
tatgtagaagt	atcccatcac	aggaagacag	agtggtcgtg	taccttatga	gccacccacg	600
tttggcaactg	aattcacagc	agtcctgtac	gaaccagcag	gtgagactta	tgaattgtcg	660
gggatgaacc	gcgctccaat	tttaactcatt	gttactctgg	aaaccagaga	tgggcaagtc	720
ctgggcgcgac	gctgctttga	ggcccgagtc	tgctgtctgc	cagggaagaga	cagggaagcg	780
gatgaagata	gcacagaaaa	gcagcaagtt	tcggacagta	caagaacagg	tgaattgtcg	840
aagcgcgcctg	ttcgtcagaa	cacacatggt	atccagatga	catccatcaa	gaacgaaga	900
tcocacagatg	atgaactgtt	atacttacca	gtgagagggc	gtgagactta	tgaattgtcg	960
ttgaagataga	aagagctcct	ggaaactcatg	cagtaccttc	ctcagcacac	aattgaaagc	1020
tacagcgcaac	agcaaacagca	gcagcacaccg	cacttacttc	agaaacagac	ctcaatacac	1080
tctccatctt	catatggtaa	cagctcccca	cctctgaaca	aaatgaacag	catgaacaag	1140
ctgcctcttg	tgagccagct	tatcaacctc	cagcagcgca	acgcctcac	tctcaaacac	1200
attctctgatg	gcattgggagc	caacattccc	atgatgggca	ccacatctac	aatgctgtga	1260
gacatgaatg	gactcagccc	caccagggca	ctccctcccc	caactctcat	gccatccacc	1320

tcccactgca cacccccacc tccgtatccc acagattgca gcattgtcag gatctggcaa 1380  
gtctga 1386

<210> 337

<211> 1551

<212> DNA

<213> Homo sapiens

<400> 337

atgtcccaga gcacacagac aaatgaatto ctccagtcag aggtttttcca gcatactcgg 60  
gattttctgg aacagcctat atgttcagtt cagccattg acttgaactt tgtggatgaa 120  
ccatcagaag atgggtcgac aaacaagatt gagattagca tggactgtat ccgcatgcag 180  
gactcggacc tgagtgaacc catgtggcca cagtacacga acctggggct cctgaacagc 240  
atggaccagc agattcagaa cggctcctcg tccaccagtc cctataacac agaccacgcg 300  
cagaacacgc tcacggcgcc ctcgcctac gcacagccca gctccacett cgtatgctctc 360  
tctcatcacc cgcgcattccc ctccaacacc gactaccag gcccgacagc ttctgacgtg 420  
tccttcacgc agtcgagcac cgccaagtgc gccacctgga cgtattccac tgaactgaag 480  
aaactctact gccaaattgc aaagacatgc cccatccaga tcaaggtgat gacctcaact 540  
cctcagggag ctgttatccg cgccatgcct gtctacaaaa aagctgagca cgtcacggag 600  
gtggtgaagc ggtgcccaa ccatgagctg agccgtgaat tcaacgaggg acagattgcc 660  
cctcctagtc atttgattcg agtagagggg aacagccatg cccagtatgt agaagatccc 720  
atcacaggaa gacagagtggt gctggtacct tatgagccac ccaggttgg cactgaattc 780  
acgacagtct tgtacaattt catgtgtaac agcagttgtg ttggagggat gaaccgcgtg 840  
caaattttaa ctattgtttac totggaaacc agagatgggc aagtcctggg ccgacgctgc 900  
ttgaggccc ggactctgtgc ttgcccagga agagacagga aggcggatga agatagcatc 960  
agaaagcagc agaatttcgg cagtacaaa aacgggtgat gtacgaaggc cccgtttcgt 1020  
cagaacacac atggtattcca gatgacatcc atcaagaaac gaagatcccc agatgatgaa 1080  
ctgttatact taccagtgaag gggccgtgag acttatgaaa tgcgttcgaa gatcaaagag 1140  
tccttgaac tcattgcagta ccttctctcag cacacaattg aaactgtacg gaacacagca 1200  
cagcagcagc accagcactt acttcagaaa cagacctcaa tacagtctcc atcttcatat 1260  
ggtaacagct ccccaactct gaacaaaatg aacagcatga acaagctgcc ttctgtgagc 1320  
cagcttatca accctcagca gcgcaacgcc ctactccta caaccattcc tgatggcatc 1380  
ggagccaaca ttcccatgat atgccaatgg ctggagacat gaatggactc 1440  
agcccccaac aggcactccc tccccactc tccatgccat ccacctccca ctgcacaccc 1500  
ccactccgt atccacacaga ttgcagcatt gtcaggatct ggcaagtctg a 1551

<210> 338

<211> 586

<212> PRT

<213> Homo sapiens

<400> 338

Met Leu Tyr Leu Glu Asn Asn Ala Gln Thr Gln Phe Ser Glu Pro Gln  
1 5 10 15  
Tyr Thr Asn Leu Gly Leu Leu Asn Ser Met Asp Gln Gln Ile Arg Asn  
20 25 30  
Gly Ser Ser Ser Thr Ser Pro Tyr Asn Thr Asp His Ala Gln Asn Ser  
35 40 45  
Val Thr Ala Pro Ser Pro Tyr Ala Gln Pro Ser Pro Thr Phe Asp Ala  
50 55 60  
Leu Ser Pro Ser Pro Ala Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro  
65 70 75 80  
His Ser Ser Asp Val Ser Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala  
85 90 95  
Thr Trp Thr Tyr Ser Thr Glu Leu Lys Lys Leu Tyr Cys Gln Ile Ala  
100 105 110  
Lys Thr Cys Pro Ile Gln Ile Lys Val Met Thr Pro Pro Gln Gly  
115 120 125

Ala Val Ile Arg Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr  
 130 135 140  
 Glu Val Val Lys Arg Cys Pro Asn His Glu Leu Ser Arg Glu Phe Asn  
 145 150 155 160  
 Glu Gly Gln Ile Ala Pro Pro Ser His Leu Ile Arg Val Glu Gly Asn  
 165 170 175  
 Ser His Ala Gln Tyr Val Glu Asp Pro Ile Thr Gly Arg Gln Ser Val  
 180 185 190  
 Leu Val Pro Tyr Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Val  
 195 200 205  
 Leu Tyr Asn Phe Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg  
 210 215 220  
 Arg Pro Ile Leu Ile Ile Val Thr Leu Glu Thr Arg Asp Gly Gln Val  
 225 230 235 240  
 Leu Gly Arg Arg Cys Phe Glu Ala Arg Ile Cys Ala Cys Pro Gly Arg  
 245 250 255  
 Asp Arg Lys Ala Asp Glu Asp Ser Ile Arg Lys Gln Gln Val Ser Asp  
 260 265 270  
 Ser Thr Lys Asn Gly Asp Gly Thr Lys Arg Pro Phe Arg Gln Asn Thr  
 275 280 285  
 His Gly Ile Gln Met Thr Ser Ile Lys Lys Arg Arg Ser Pro Asp Asp  
 290 295 300  
 Glu Leu Leu Tyr Leu Pro Val Arg Gly Arg Glu Thr Tyr Glu Met Leu  
 305 310 315 320  
 Leu Lys Ile Lys Glu Ser Leu Glu Leu Met Gln Tyr Leu Pro Gln His  
 325 330 335  
 Thr Ile Glu Thr Tyr Arg Gln Gln Gln Gln Gln His Gln His Leu  
 340 345 350  
 Leu Gln Lys Gln Thr Ser Ile Gln Ser Pro Ser Ser Tyr Gly Asn Ser  
 355 360 365  
 Ser Pro Pro Leu Asn Lys Met Asn Ser Met Asn Lys Leu Pro Ser Val  
 370 375 380  
 Ser Gln Leu Ile Asn Pro Gln Gln Arg Asn Ala Leu Thr Pro Thr Thr  
 385 390 395 400  
 Ile Pro Asp Gly Met Gly Ala Asn Ile Pro Met Met Gly Thr His Met  
 405 410 415  
 Pro Met Ala Gly Asp Met Asn Gly Leu Ser Pro Thr Gln Ala Leu Pro  
 420 425 430  
 Pro Pro Leu Ser Met Pro Ser Thr Ser His Cys Thr Pro Pro Pro Pro  
 435 440 445  
 Tyr Pro Thr Asp Cys Ser Ile Val Ser Phe Leu Ala Arg Leu Gly Cys  
 450 455 460  
 Ser Ser Cys Leu Asp Tyr Phe Thr Thr Gln Gly Leu Thr Thr Ile Tyr  
 465 470 475 480  
 Gln Ile Glu His Tyr Ser Met Asp Asp Leu Ala Ser Leu Lys Ile Pro  
 485 490 495  
 Glu Gln Phe Arg His Ala Ile Trp Lys Gly Ile Leu Asp His Arg Gln  
 500 505 510  
 Leu His Glu Phe Ser Ser Pro Ser His Leu Leu Arg Thr Pro Ser Ser  
 515 520 525  
 Ala Ser Thr Val Ser Val Gly Ser Ser Glu Thr Arg Gly Glu Arg Val  
 530 535 540  
 Ile Asp Ala Val Arg Phe Thr Leu Arg Gln Thr Ile Ser Phe Pro Pro  
 545 550 555 560  
 Arg Asp Glu Trp Asn Asp Phe Asn Phe Asp Met Asp Ala Arg Arg Asn  
 565 570 575  
 Lys Gln Gln Arg Ile Lys Glu Glu Gly Glu  
 580 585

<210> 339  
 <211> 641  
 <212> PRT  
 <213> Homo sapiens

<400> 339

```

Met Ser Gln Ser Thr Gln Thr Asn Glu Phe Leu Ser Pro Glu Val Phe
 1                               5 10 15
Gln His Ile Trp Asp Phe Leu Glu Gln Pro Ile Cys Ser Val Gln Pro
 20 25 30
Ile Asp Leu Asn Phe Val Asp Glu Pro Ser Glu Asp Gly Ala Thr Asn
 35 40 45
Lys Ile Glu Ile Ser Met Asp Cys Ile Arg Met Gln Asp Ser Asp Leu
 50 55 60
Ser Asp Pro Met Trp Pro Gln Tyr Thr Asn Leu Gly Leu Leu Asn Ser
 65 70 75 80
Met Asp Gln Gln Ile Gln Asn Gly Ser Ser Thr Ser Pro Tyr Asn
 85 90 95
Thr Asp His Ala Gln Asn Ser Val Thr Ala Pro Ser Pro Tyr Ala Gln
 100 105 110
Pro Ser Ser Thr Phe Asp Ala Leu Ser Pro Ser Pro Ala Ile Pro Ser
 115 120 125
Asn Thr Asp Tyr Pro Gly Pro His Ser Phe Asp Val Ser Phe Gln Gln
 130 135 140
Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr Ser Thr Glu Leu Lys
 145 150 155 160
Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro Ile Gln Ile Lys Val
 165 170 175
Met Thr Pro Pro Pro Gln Gly Ala Val Ile Arg Ala Met Pro Val Tyr
 180 185 190
Lys Lys Ala Glu His Val Thr Glu Val Val Lys Arg Cys Pro Asn His
 195 200 205
Glu Leu Ser Arg Glu Phe Asn Glu Gly Gln Ile Ala Pro Pro Ser His
 210 215 220
Leu Ile Arg Val Glu Gly Asn Ser His Ala Gln Tyr Val Glu Asp Pro
 225 230 235 240
Ile Thr Gly Arg Gln Ser Val Leu Val Pro Tyr Glu Pro Pro Gln Val
 245 250 255
Gly Thr Glu Phe Thr Thr Val Leu Tyr Asn Phe Met Cys Asn Ser Ser
 260 265 270
Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu Ile Ile Val Thr Leu
 275 280 285
Glu Thr Arg Asp Gly Gln Val Leu Gly Arg Arg Cys Phe Glu Ala Arg
 290 295 300
Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala Asp Glu Asp Ser Ile
 305 310 315 320
Arg Lys Gln Gln Val Ser Asp Ser Thr Lys Asn Gly Asp Gly Thr Lys
 325 330 335
Arg Pro Phe Arg Gln Asn Thr His Gly Ile Gln Met Thr Ser Ile Lys
 340 345 350
Lys Arg Arg Ser Pro Asp Asp Glu Leu Leu Tyr Leu Pro Val Arg Gly
 355 360 365
Arg Glu Thr Tyr Glu Met Leu Leu Lys Ile Lys Glu Ser Leu Glu Leu
 370 375 380
Met Gln Tyr Leu Pro Gln His Thr Ile Glu Thr Tyr Arg Gln Gln Gln
 385 390 395 400

```

Gln Gln Gln His Gln His Leu Leu Gln Lys Gln Thr Ser Ile Gln Ser  
 405 410 415  
 Pro Ser Ser Tyr Gly Asn Ser Ser Pro Leu Asn Lys Met Asn Ser  
 420 425 430  
 Met Asn Lys Leu Pro Ser Val Ser Gln Leu Ile Asn Pro Gln Gln Arg  
 435 440 445  
 Asn Ala Leu Thr Pro Thr Thr Ile Pro Asp Gly Met Gly Ala Asn Ile  
 450 455 460  
 Pro Met Met Gly Thr His Met Pro Met Ala Gly Asp Met Asn Gly Leu  
 465 470 475 480  
 Ser Pro Thr Gln Ala Leu Pro Pro Pro Leu Ser Met Pro Ser Thr Ser  
 485 490 495  
 His Cys Thr Pro Pro Pro Tyr Pro Thr Asp Cys Ser Ile Val Gly  
 500 505 510  
 Phe Leu Ala Arg Leu Gly Cys Ser Ser Cys Leu Asp Tyr Phe Thr Thr  
 515 520 525  
 Gln Gly Leu Thr Thr Ile Tyr Gln Ile Glu His Tyr Ser Met Asp Asp  
 530 535 540  
 Leu Ala Ser Leu Lys Ile Pro Glu Gln Phe Arg His Ala Ile Trp Lys  
 545 550 555 560  
 Gly Ile Leu Asp His Arg Gln Leu His Glu Phe Ser Ser Pro Ser His  
 565 570 575  
 Leu Leu Arg Thr Pro Ser Ser Ala Ser Thr Val Ser Val Gly Ser Ser  
 580 585 590  
 Glu Thr Arg Gly Glu Arg Val Ile Asp Ala Val Arg Phe Thr Leu Arg  
 595 600 605  
 Gln Thr Ile Ser Phe Pro Pro Arg Asp Glu Trp Asn Asp Phe Asn Phe  
 610 615 620  
 Asp Met Asp Ala Arg Arg Asn Lys Gln Gln Arg Ile Lys Glu Glu Gly  
 625 630 635 640  
 Glu

&lt;210&gt; 340

&lt;211&gt; 448

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 340

Met Ser Gln Ser Thr Gln Thr Asn Glu Phe Leu Ser Pro Glu Val Phe  
 1 5 10 15  
 Gln His Ile Trp Asp Phe Leu Glu Gln Pro Ile Cys Ser Val Gln Pro  
 20 25 30  
 Ile Asp Leu Asn Phe Val Asp Glu Pro Ser Glu Asp Gly Ala Thr Asn  
 35 40 45  
 Lys Ile Glu Ile Ser Met Asp Cys Ile Arg Met Gln Asp Ser Asp Leu  
 50 55 60  
 Ser Asp Pro Met Trp Pro Gln Tyr Thr Asn Leu Gly Leu Leu Asn Ser  
 65 70 75 80  
 Met Asp Gln Gln Ile Gln Asn Gly Ser Ser Ser Thr Ser Pro Tyr Asn  
 85 90 95  
 Thr Asp His Ala Gln Asn Ser Val Thr Ala Pro Ser Pro Tyr Ala Gln  
 100 105 110  
 Pro Ser Ser Thr Phe Asp Ala Leu Ser Pro Ser Pro Ala Ile Pro Ser  
 115 120 125  
 Asn Thr Asp Tyr Pro Gly Pro His Ser Phe Asp Val Ser Phe Gln Gln  
 130 135 140

Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr Ser Thr Glu Leu Lys  
 145 150 155 160  
 Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro Ile Gln Ile Lys Val  
 165 170 175  
 Met Thr Pro Pro Pro Gln Gly Ala Val Ile Arg Ala Met Pro Val Tyr  
 180 185 190  
 Lys Lys Ala Glu His Val Thr Glu Val Val Lys Arg Cys Pro Asn His  
 195 200 205  
 Glu Leu Ser Arg Glu Phe Asn Glu Gly Gln Ile Ala Pro Pro Ser His  
 210 215 220  
 Leu Ile Arg Val Glu Gly Asn Ser His Ala Gln Tyr Val Glu Asp Pro  
 225 230 235 240  
 Ile Thr Gly Arg Gln Ser Val Leu Val Pro Tyr Glu Pro Pro Gln Val  
 245 250 255  
 Gly Thr Glu Phe Thr Thr Val Leu Tyr Asn Phe Met Cys Asn Ser Ser  
 260 265 270  
 Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu Ile Ile Val Thr Leu  
 275 280 285  
 Glu Thr Arg Asp Gly Gln Val Leu Gly Arg Arg Cys Phe Glu Ala Arg  
 290 295 300  
 Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala Asp Glu Asp Ser Ile  
 305 310 315 320  
 Arg Lys Gln Gln Val Ser Asp Ser Thr Lys Asn Gly Asp Gly Thr Lys  
 325 330 335  
 Arg Pro Phe Arg Gln Asn Thr His Gly Ile Gln Met Thr Ser Ile Lys  
 340 345 350  
 Lys Arg Arg Ser Pro Asp Asp Glu Leu Leu Tyr Leu Pro Val Arg Gly  
 355 360 365  
 Arg Glu Thr Tyr Glu Met Leu Leu Lys Ile Lys Glu Ser Leu Glu Leu  
 370 375 380  
 Met Gln Tyr Leu Pro Gln His Thr Ile Glu Thr Tyr Arg Gln Gln Gln  
 385 390 395 400  
 Gln Gln Gln His Gln His Leu Leu Gln Lys His Leu Leu Ser Ala Cys  
 405 410 415  
 Phe Arg Asn Glu Leu Val Glu Pro Arg Arg Glu Thr Pro Lys Gln Ser  
 420 425 430  
 Asp Val Phe Phe Arg His Ser Lys Pro Pro Asn Arg Ser Val Tyr Pro  
 435 440 445

&lt;210&gt; 341

&lt;211&gt; 356

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 341

Met Leu Tyr Leu Glu Asn Asn Ala Gln Thr Gln Phe Ser Glu Pro Gln  
 1 5 10 15  
 Tyr Thr Asn Leu Gly Leu Leu Asn Ser Met Asp Gln Gln Ile Gln Asn  
 20 25 30  
 Gly Ser Ser Ser Thr Ser Pro Tyr Asn Thr Asp His Ala Gln Asn Ser  
 35 40 45  
 Val Thr Ala Pro Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Ala  
 50 55 60  
 Leu Ser Pro Ser Pro Ala Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro  
 65 70 75 80  
 His Ser Phe Asp Val Ser Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala  
 85 90 95

```

Thr Trp Thr Tyr Ser Thr Glu Leu Lys Lys Leu Tyr Cys Gln Ile Ala
      100      105      110
Lys Thr Cys Pro Ile Gln Ile Lys Val Met Thr Pro Pro Gln Gly
      115      120      125
Ala Val Ile Arg Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr
      130      135      140
Glu Val Val Lys Arg Cys Pro Asn His Glu Leu Ser Arg Glu Phe Asn
145      150      155      160
Glu Gly Gln Ile Ala Pro Pro Ser His Leu Ile Arg Val Glu Gly Asn
      165      170      175
Ser His Ala Gln Tyr Val Glu Asp Pro Ile Thr Gly Arg Gln Ser Val
      180      185      190
Leu Val Pro Tyr Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Val
      195      200      205
Leu Tyr Asn Phe Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg
      210      215      220
Arg Pro Ile Leu Ile Ile Val Thr Leu Glu Thr Arg Asp Gly Gln Val
225      230      235      240
Leu Gly Arg Arg Cys Phe Glu Ala Arg Ile Cys Ala Cys Pro Gly Arg
      245      250      255
Asp Arg Lys Ala Asp Glu Asp Ser Ile Arg Lys Gln Gln Val Ser Asp
      260      265      270
Ser Thr Lys Asn Gly Asp Gly Thr Lys Arg Pro Ser Arg Gln Asn Thr
      275      280      285
His Gly Ile Gln Met Thr Ser Ile Lys Lys Arg Arg Ser Pro Asp Asp
290      295      300
Glu Leu Leu Tyr Leu Pro Val Arg Gly Arg Glu Thr Tyr Glu Met Leu
305      310      315      320
Leu Lys Ile Lys Glu Ser Leu Glu Leu Met Gln Tyr Leu Pro Gln His
      325      330      335
Thr Ile Glu Thr Tyr Arg Gln Gln Gln Gln Gln His Gln His Leu
      340      345      350
Leu Gln Lys Gln
      355

```

&lt;210&gt; 342

&lt;211&gt; 680

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 342

```

Met Asn Phe Glu Thr Ser Arg Cys Ala Thr Leu Gln Tyr Cys Pro Asp
 1      5      10      15
Pro Tyr Ile Gln Arg Phe Val Glu Thr Pro Ala His Phe Ser Trp Lys
      20      25      30
Glu Ser Tyr Tyr Arg Ser Thr Met Ser Gln Ser Thr Gln Thr Asn Glu
      35      40      45
Phe Leu Ser Pro Glu Val Phe Gln His Ile Trp Asp Phe Leu Glu Gln
50      55      60
Pro Ile Cys Ser Val Gln Pro Ile Asp Leu Asn Phe Val Asp Glu Pro
65      70      75      80
Ser Glu Asp Gly Ala Thr Asn Lys Ile Glu Ile Ser Met Asp Cys Ile
      85      90      95
Arg Met Gln Asp Ser Asp Leu Ser Asp Pro Met Trp Pro Gln Tyr Thr
      100      105      110
Asn Leu Gly Leu Leu Asn Ser Met Asp Gln Gln Ile Gln Asn Gly Ser
      115      120      125

```



Ser Ser Thr Ser Pro Tyr Asn Thr Asp His Ala Gln Asn Ser Val Thr  
 130 135 140  
 Ala Pro Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Ala Leu Ser  
 145 150 155 160  
 Pro Ser Pro Ala Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro His Ser  
 165 170 175  
 Phe Asp Val Ser Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala Thr Trp  
 180 185 190  
 Thr Tyr Ser Thr Glu Leu Lys Lys Leu Tyr Cys Gln Ile Ala Lys Thr  
 195 200 205  
 Cys Pro Ile Gln Ile Lys Val Met Thr Pro Pro Pro Gln Gly Ala Val  
 210 215 220  
 Ile Arg Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr Glu Val  
 225 230 235 240  
 Val Lys Arg Cys Pro Asn His Glu Leu Ser Arg Glu Phe Asn Glu Gly  
 245 250 255  
 Gln Ile Ala Pro Pro Ser His Leu Ile Arg Val Glu Gly Asn Ser His  
 260 265 270  
 Ala Gln Tyr Val Glu Asp Pro Ile Thr Gly Arg Gln Ser Val Leu Val  
 275 280 285  
 Pro Tyr Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Val Leu Tyr  
 290 295 300  
 Asn Phe Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg Arg Pro  
 305 310 315 320  
 Ile Leu Ile Ile Val Thr Leu Glu Thr Arg Asp Gly Gln Val Leu Gly  
 325 330 335  
 Arg Arg Cys Phe Glu Ala Arg Ile Cys Ala Cys Pro Gly Arg Asp Arg  
 340 345 350  
 Lys Ala Asp Glu Asp Ser Ile Arg Lys Gln Gln Val Ser Asp Ser Thr  
 355 360 365  
 Lys Asn Gly Asp Gly Thr Lys Arg Pro Phe Arg Gln Asn Thr His Gly  
 370 375 380  
 Ile Gln Met Thr Ser Ile Lys Lys Arg Arg Ser Pro Asp Asp Glu Leu  
 385 390 395 400  
 Leu Tyr Leu Pro Val Arg Gly Arg Glu Thr Tyr Glu Met Leu Leu Lys  
 405 410 415  
 Ile Lys Glu Ser Leu Glu Leu Met Gln Tyr Leu Pro Gln His Thr Ile  
 420 425 430  
 Glu Thr Tyr Arg Gln Gln Gln Gln Gln His Gln His Leu Leu Gln  
 435 440 445  
 Lys Gln Thr Ser Ile Gln Ser Pro Ser Ser Tyr Gly Asn Ser Ser Pro  
 450 455 460  
 Pro Leu Asn Lys Met Asn Ser Met Asn Lys Leu Pro Ser Val Ser Gln  
 465 470 475 480  
 Leu Ile Asn Pro Gln Gln Arg Asn Ala Leu Thr Pro Thr Thr Ile Pro  
 485 490 495  
 Asp Gly Met Gly Ala Asn Ile Pro Met Met Gly Thr His Met Pro Met  
 500 505 510  
 Ala Gly Asp Met Asn Gly Leu Ser Pro Thr Gln Ala Leu Pro Pro Pro  
 515 520 525  
 Leu Ser Met Pro Ser Thr Ser Gln Cys Thr Pro Pro Pro Tyr Pro  
 530 535 540  
 Thr Asp Cys Ser Ile Val Ser Phe Leu Ala Arg Leu Gly Cys Ser Ser  
 545 550 555 560  
 Cys Leu Asp Tyr Phe Thr Thr Gln Gly Leu Thr Thr Ile Tyr Gln Ile  
 565 570 575  
 Glu His Tyr Ser Met Asp Asp Leu Ala Ser Leu Lys Ile Pro Glu Gln  
 580 585 590

150

Phe Arg His Ala Ile Trp Lys Gly Ile Leu Asp His Arg Gln Leu His  
 595 600 605  
 Glu Phe Ser Ser Pro Ser His Leu Leu Arg Thr Pro Ser Ser Ala Ser  
 610 615 620  
 Thr Val Ser Val Gly Ser Ser Glu Thr Arg Gly Glu Arg Val Ile Asp  
 625 630 635 640  
 Ala Val Arg Phe Thr Leu Arg Gln Thr Ile Ser Phe Pro Pro Arg Asp  
 645 650 655  
 Glu Trp Asn Asp Phe Asn Phe Asp Met Asp Ala Arg Arg Asn Lys Gln  
 660 665 670  
 Gln Arg Ile Lys Glu Glu Gly Glu  
 675 680

&lt;210&gt; 343

&lt;211&gt; 461

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 343

Met Leu Tyr Leu Glu Asn Asn Ala Gln Thr Gln Phe Ser Glu Pro Gln  
 1 5 10 15  
 Tyr Thr Asn Leu Gly Leu Leu Asn Ser Met Asp Gln Gln Ile Gln Asn  
 20 25 30  
 Gly Ser Ser Ser Thr Ser Pro Tyr Asn Thr Asp His Ala Gln Asn Ser  
 35 40 45  
 Val Thr Ala Pro Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Ala  
 50 55 60  
 Leu Ser Pro Ser Pro Ala Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro  
 65 70 75 80  
 His Ser Phe Asp Val Ser Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala  
 85 90 95  
 Thr Trp Thr Tyr Ser Thr Glu Leu Lys Lys Leu Tyr Cys Gln Ile Ala  
 100 105 110  
 Lys Thr Cys Pro Ile Gln Ile Lys Val Met Thr Pro Pro Gln Gln Gly  
 115 120 125  
 Ala Val Ile Arg Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr  
 130 135 140  
 Glu Val Val Lys Arg Cys Pro Asn His Glu Leu Ser Arg Glu Phe Asn  
 145 150 155 160  
 Glu Gly Gln Ile Ala Pro Pro Ser His Leu Ile Arg Val Glu Gly Asn  
 165 170 175  
 Ser His Ala Gln Tyr Val Glu Asp Pro Ile Thr Gly Arg Gln Ser Val  
 180 185 190  
 Leu Val Pro Tyr Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Val  
 195 200 205  
 Leu Tyr Asn Phe Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg  
 210 215 220  
 Arg Pro Ile Leu Ile Ile Val Thr Leu Glu Thr Arg Asp Gly Gln Val  
 225 230 235 240  
 Leu Gly Arg Arg Cys Phe Glu Ala Arg Ile Cys Ala Cys Pro Gly Arg  
 245 250 255  
 Asp Arg Lys Ala Asp Glu Asp Ser Ile Arg Lys Gln Gln Val Ser Asp  
 260 265 270  
 Ser Thr Lys Asn Gly Asp Gly Thr Lys Arg Pro Phe Arg Gln Asn Thr  
 275 280 285  
 His Gly Ile Gln Met Thr Ser Ile Lys Lys Arg Arg Ser Pro Asp Asp  
 290 295 300

Glu Leu Leu Tyr Leu Pro Val Arg Gly Arg Glu Thr Tyr Glu Met Leu  
 305 310 315 320  
 Leu Lys Ile Lys Glu Ser Leu Glu Leu Met Gln Tyr Leu Pro Gln His  
 325 330 335  
 Thr Ile Glu Thr Tyr Arg Gln Gln Gln Gln His Gln His Leu  
 340 345 350  
 Leu Gln Lys Gln Thr Ser Ile Gln Ser Pro Ser Ser Tyr Gly Asn Ser  
 355 360 365  
 Ser Pro Pro Leu Asn Lys Met Asn Ser Met Asn Lys Leu Pro Ser Val  
 370 375 380  
 Ser Gln Leu Ile Asn Pro Gln Gln Arg Asn Ala Leu Thr Pro Thr Thr  
 385 390 395 400  
 Ile Pro Asp Gly Met Gly Ala Asn Ile Pro Met Met Gly Thr His Met  
 405 410 415  
 Pro Met Ala Gly Asp Met Asn Gly Leu Ser Pro Thr Gln Ala Leu Pro  
 420 425 430  
 Pro Pro Leu Ser Met Pro Ser Thr Ser His Cys Thr Pro Pro Pro  
 435 440 445  
 Tyr Pro Thr Asp Cys Ser Ile Val Arg Ile Trp Gln Val  
 450 455 460

&lt;210&gt; 344

&lt;211&gt; 516

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 344

Met Ser Gln Ser Thr Gln Thr Asn Glu Phe Leu Ser Pro Glu Val Phe  
 1 5 10 15  
 Gln His Ile Trp Asp Phe Leu Glu Gln Pro Ile Cys Ser Val Gln Pro  
 20 25 30  
 Ile Asp Leu Asn Phe Val Asp Glu Pro Ser Glu Asp Gly Ala Thr Asn  
 35 40 45  
 Lys Ile Glu Ile Ser Met Asp Cys Ile Arg Met Gln Asp Ser Asp Leu  
 50 55 60  
 Ser Asp Pro Met Trp Pro Gln Tyr Thr Asn Leu Gly Leu Leu Asn Ser  
 65 70 75 80  
 Met Asp Gln Gln Ile Gln Asn Gly Ser Ser Ser Thr Ser Pro Tyr Asn  
 85 90 95  
 Thr Asp His Ala Gln Asn Ser Val Thr Ala Pro Ser Pro Tyr Ala Gln  
 100 105 110  
 Pro Ser Ser Thr Phe Asp Ala Leu Ser Pro Ser Pro Ala Ile Pro Ser  
 115 120 125  
 Asn Thr Asp Tyr Pro Gly Pro His Ser Phe Asp Val Ser Phe Gln Gln  
 130 135 140  
 Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr Ser Thr Glu Leu Lys  
 145 150 155 160  
 Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro Ile Gln Ile Lys Val  
 165 170 175  
 Met Thr Pro Pro Gln Gly Ala Val Ile Arg Ala Met Pro Val Tyr  
 180 185 190  
 Lys Lys Ala Glu His Val Thr Glu Val Val Lys Arg Cys Pro Asn His  
 195 200 205  
 Glu Leu Ser Arg Glu Phe Asn Glu Gly Gln Ile Ala Pro Pro Ser His  
 210 215 220  
 Leu Ile Arg Val Glu Gly Asn Ser His Ala Gln Tyr Val Glu Asp Pro  
 225 230 235 240

Ile Thr Gly Arg Gln Ser Val Leu Val Pro Tyr Glu Pro Pro Gln Val  
 245 250 255  
 Gly Thr Glu Phe Thr Thr Val Leu Tyr Asn Phe Met Cys Asn Ser Ser  
 260 265 270  
 Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu Ile Ile Val Thr Leu  
 275 280 285  
 Glu Thr Arg Asp Gly Gln Val Leu Gly Arg Arg Cys Phe Glu Ala Arg  
 290 295 300  
 Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala Asp Glu Asp Ser Ile  
 305 310 315 320  
 Arg Lys Gln Gln Val Ser Asp Ser Thr Lys Asn Gly Asp Gly Thr Lys  
 325 330 335  
 Arg Pro Phe Arg Gln Asn Thr His Gly Ile Gln Met Thr Ser Ile Lys  
 340 345 350  
 Lys Arg Arg Ser Pro Asp Asp Glu Leu Leu Tyr Leu Pro Val Arg Gly  
 355 360 365  
 Arg Glu Thr Tyr Glu Met Leu Leu Lys Ile Lys Glu Ser Leu Glu Leu  
 370 375 380  
 Met Gln Tyr Leu Pro Gln His Thr Ile Glu Thr Tyr Arg Gln Gln Gln  
 385 390 395 400  
 Gln Gln Gln His Gln His Leu Leu Gln Lys Gln Thr Ser Ile Gln Ser  
 405 410 415  
 Pro Ser Ser Tyr Gly Asn Ser Ser Pro Leu Asn Lys Met Asn Ser  
 420 425 430  
 Met Asn Lys Leu Pro Ser Val Ser Gln Leu Ile Asn Pro Gln Gln Arg  
 435 440 445  
 Asn Ala Leu Thr Pro Thr Thr Ile Pro Asp Gly Met Gly Ala Asn Ile  
 450 455 460  
 Pro Met Met Gly Thr His Met Pro Met Ala Gly Asp Met Asn Gly Leu  
 465 470 475 480  
 Ser Pro Thr Gln Ala Leu Pro Pro Pro Leu Ser Met Pro Ser Thr Ser  
 485 490 495  
 His Cys Thr Pro Pro Pro Pro Tyr Pro Thr Asp Cys Ser Ile Val Arg  
 500 505 510  
 Ile Trp Gln Val  
 515

&lt;210&gt; 345

&lt;211&gt; 1800

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 345

ggcgcctcatt gccactgcag tgaactaaagc tgggaagacg ctgggtcagtt cacctgcccc 60  
 actggttggtt ttttaaacaa attctgatac aggcgacatc ctcaactgacc gagcaaaagt 120  
 tgacattcgt atcatcacg tgcaccattg gcttctaggc actccagtggt ggtgagagaa 180  
 ggaggtctga aacctcgca gagggatctt gccctcattc ttgggtctgt aaacactggc 240  
 agtcgttggg aacaggactc agggataaac cagcgcaatg gattggggga cgctgcacac 300  
 ttctcatcggg ggtgtcaaca aacactccac cagcatcggg aaggtgtgga tcacagtcat 360  
 ctttattttc cgagtcatga tcctagtgtt ggctgcccag gaagtgtggg gtgacagaca 420  
 agaggacttc gtctgcaaca cactgcaacc gggatgcaaa aatgtgtgct atgaccactt 480  
 tttccgggtg tccacatccc ggcgtgtggc cctccagctg atcttcgtot ccaccccagc 540  
 gctgctgggt gccatgcatg tggcctaacta caggcacgaa accactcgca agttcaggcg 600  
 aggagagaag aggaatgatt tcaaaagacat agaggacatt aaaaagcaca aggttcggat 660  
 agaggggtat cgttgggtgga cgtacaccag cagcatcttt ttccgaatca agtttgaagc 720  
 agcctttatg tatgtgtttt acttccttta caatgggttac cacctgcocct ggtgtgtgaa 780  
 atgtgggatt gacccctgcc ccaaccttgt tgactgcttt atttctaggc caacagagaa 840

```

gaccgtgttt accattttta tgatttctgc gtctgtgatt tgcatgctgc ttaacgtggc 900
agagttgtgc tacctgtgtc tgaagtgtg ttttaggaga tcaagagag cagacagcga 960
aaaaaatcac cccaatcatg cccataaagg gagtaagcag aatgaaatga atgagctgat 1020
ttcagatagt ggtcaaaatg caatcacagc ttcccaagc taacacattc aaggtaaaat 1080
gtagctgcgt cataaggaga ctctgtgtct ctccagaagg caataccaac ctgaaagtgc 1140
ctctgtgtag ctgaagagtt tgtaaattgac ttccataata aatagacact tgagttaact 1200
ttttgttaga tacttgcctc attcatacac aacgtaatca aatatgtggt ccattctcta 1260
aaacaagaga ctgcttgaca aaggagcatt gcagtcactt tgacaggttc cttttaagtg 1320
gactctctga caaagtgggt actttctgaa aatttatata actgtgtgtg ataaggaaac 1380
ttattccagg aattgatacg tttatttaga aaagatattt ttataggctt ggatgttttt 1440
agttccgact ttgaatttat ataaagtatt tttataatga ctggctctcc ttacctggaa 1500
aaacatgcga tgtagttttt agaattacac cacaagtatc taaatttcca acttacaaga 1560
ggctctatct tgtaaattatt gttttgcatt gtctgttggc aaatttggta actgtcatga 1620
tacgcttaag gtgggaaagt gtccattgca caatatattt ttactgcttt ctgaatgtag 1680
acggaacagt gtggaagcag aaggcttttt taactcatcc gtttggccga tcgttgacga 1740
ccactgggag atgtggatgt ggttgccctc ttttgctcgt cccgtgggct taacctctct 1800

```

&lt;210&gt; 346

&lt;211&gt; 261

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 346

```

Met Asp Trp Gly Thr Leu His Thr Phe Ile Gly Gly Val Asn Lys His
  1           5           10           15
Ser Thr Ser Ile Gly Lys Val Trp Ile Thr Val Ile Phe Ile Phe Arg
      20           25           30
Val Met Ile Leu Val Val Ala Ala Gln Glu Val Trp Gly Asp Glu Gln
      35           40           45
Glu Asp Phe Val Cys Asn Thr Leu Gln Pro Gly Cys Lys Asn Val Cys
      50           55           60
Tyr Asp His Phe Phe Pro Val Ser His Ile Arg Leu Trp Ala Leu Gln
      65           70           75           80
Leu Ile Phe Val Ser Thr Pro Ala Leu Leu Val Ala Met His Val Ala
      85           90           95
Tyr Tyr Arg His Glu Thr Thr Arg Lys Phe Arg Arg Gly Glu Lys Arg
      100          105          110
Asn Asp Phe Lys Asp Ile Glu Asp Ile Lys Lys His Lys Val Arg Ile
      115          120          125
Glu Gly Ser Leu Trp Trp Thr Tyr Thr Ser Ser Ile Phe Phe Arg Ile
      130          135          140
Ile Phe Glu Ala Ala Phe Met Tyr Val Phe Tyr Phe Leu Tyr Asn Gly
      145          150          155          160
Tyr His Leu Pro Trp Val Leu Lys Cys Gly Ile Asp Pro Cys Pro Asn
      165          170          175
Leu Val Asp Cys Phe Ile Ser Arg Pro Thr Glu Lys Thr Val Phe Thr
      180          185          190
Ile Phe Met Ile Ser Ala Ser Val Ile Cys Met Leu Leu Asn Val Ala
      195          200          205
Glu Leu Cys Tyr Leu Leu Leu Lys Val Cys Phe Arg Arg Ser Lys Arg
      210          215          220
Ala Gln Thr Gln Lys Asn His Pro Asn His Ala Leu Lys Glu Ser Lys
      225          230          235          240
Gln Asn Glu Met Asn Glu Leu Ile Ser Asp Ser Gly Gln Asn Ala Ile
      245          250          255
Thr Gly Phe Pro Ser
      260

```

<210> 347  
 <211> 1740  
 <212> DNA  
 <213> Homo sapiens

<400> 347  
 atgaacaaac tgtatatcgg aaacctcagc gagaacgccc cccctcggga cctagaaagt 60  
 atcttcaagg acgccaagat cccgggtgtcg ggacccttcc tgggtgaagac tgggtacgcy 120  
 ttogtggact gcccggaagc gagctggggc ctcaaggcca tcgaggcgctt ttacaggtaa 180  
 atagaactgc acgggaaccc catagaagtt gagcactcgg tcccaaaaag gcaaggatt 240  
 cggaaacttc agatacagaa tatcccgccct catttacagt gggagggtgct ggatagttta 300  
 ctagtccagt atggagtggt ggagagctgt gagcaagtga acactgaactc ggaactgca 360  
 gttgtaaatg taacctattc cagtaaggac caagctagac aagcactaga caaactgaat 420  
 ggatttcagt tagagaattt cacottgaaa gttagcctata tccctgatga aacggccgcc 480  
 cagcaaaaac ccttgcagca gccccgaggt cgcggggggc ttgggcagag gggctctcca 540  
 aggcagggggt ctccaggatc cgtatccaag cagaaacccat gtgatttgcc tctgcgctg 600  
 ctggttccca cccaatttgt tggagcccat ataggaaaag aaggtgccac cattcggaac 660  
 atcaccaaac agaccagtc taaaatcgat gtcccacgta aagaaaatgc gggggctgct 720  
 gagaagtgcga ttactatcct ctctactcct gaaggcactc ctgcggcttg taagcttatt 780  
 ctggagatta tgcataagga agctcaagat ataaaattca cagaagagat ccccttgaag 840  
 attttagctc ataataactt tgttggagct cttatttgga aagaaggag aaactcttaa 900  
 aaaattgagc aagacacaga cactaaaato acgatattcc cattgcagga attgacgctg 960  
 tataatccag aacgcactat tacagttaaa ggcaatgttg agacatgtgc caaagctgag 1020  
 gaggagatca tgaagaaatc caggggagctc tatgaaaatg atattgtctc tatgaattct 1080  
 caagcacatt taattctctg attaaaatctc aacgccttgg gtctgttccc accoactcoa 1140  
 gggatgccac ctcccacctc agggcccccct tcagccatga ctctcccta cccgagttt 1200  
 gacgaactcag aaacggagac tgttcatctg tttatccagc ctctatcagt cggtgccatc 1260  
 atcgccaagc agggccaagc catcaagcag cttctctgct ttgctggagc ttcaattaa 1320  
 attgctccag cggaagcacc agatgctaaa gtgaggatgg tgattatcac tggaccacca 1380  
 gaggctcagt tcaaggctca gggagaagatt tatggaaaaa ttaagaaga aaactttgtt 1440  
 agtcctaagt aagaggtgaa acttgaagct catatcagag tgccatcctt tgctgtcggc 1500  
 agagttattg gaaaaggagg caaaacgggtc aatgaacttc agaatttgctc aagtgcagaa 1560  
 gttgtgtctc ctctgtacca gacacctgat gagaatgacc aagtgggttg caaaataact 1620  
 ggtaacttct atgcttgcca ggttgcccag agaaaaattc aggaaattct gactcaggt 1680  
 aagcagcacc aacaacagaa ggtctctgcaa agtggaccac ctacgtcaag acggaagtaa 1740

<210> 348  
 <211> 579  
 <212> PRT  
 <213> Homo sapiens

<400> 348  
 Met Asn Lys Leu Tyr Ile Gly Asn Leu Ser Glu Asn Ala Ala Pro Ser  
 1 5 10 15  
 Asp Leu Glu Ser Ile Phe Lys Asp Ala Lys Ile Pro Val Ser Gly Pro  
 20 25 30  
 Phe Leu Val Lys Thr Gly Tyr Ala Phe Val Asp Cys Pro Asp Glu Ser  
 35 40 45  
 Trp Ala Leu Lys Ala Ile Glu Ala Leu Ser Gly Lys Ile Glu Leu His  
 50 55 60  
 Gly Lys Pro Ile Glu Val Glu His Ser Val Pro Lys Arg Gln Arg Ile  
 65 70 75 80  
 Arg Lys Leu Gln Ile Arg Asn Ile Pro Pro His Leu Gln Trp Glu Val  
 85 90 95  
 Leu Asp Ser Leu Leu Val Gln Tyr Gly Val Val Glu Ser Cys Glu Gln

100				105				110							
Val	Asn	Thr	Asp	Ser	Glu	Thr	Ala	Val	Val	Asn	Val	Thr	Tyr	Ser	Ser
115				120				125							
Lys	Asp	Gln	Ala	Arg	Gln	Ala	Leu	Asp	Lys	Leu	Asn	Gly	Phe	Gln	Leu
130				135				140							
Glu	Asn	Phe	Thr	Leu	Lys	Val	Ala	Tyr	Ile	Pro	Asp	Glu	Thr	Ala	Ala
145				150				155							
Gln	Gln	Asn	Pro	Leu	Gln	Gln	Pro	Arg	Gly	Arg	Arg	Gly	Leu	Gly	160
165				170				175							
Arg	Gly	Ser	Ser	Arg	Gln	Gly	Ser	Pro	Gly	Ser	Val	Ser	Lys	Gln	Lys
180				185				190							
Pro	Cys	Asp	Leu	Pro	Leu	Arg	Leu	Val	Pro	Thr	Gln	Phe	Val	Gly	
195				200				205							
Ala	Ile	Ile	Gly	Lys	Glu	Gly	Ala	Thr	Ile	Arg	Asn	Ile	Thr	Lys	Gln
210				215				220							
Thr	Gln	Ser	Lys	Ile	Asp	Val	His	Arg	Lys	Glu	Asn	Ala	Gly	Ala	Ala
225				230				235							
Glu	Lys	Ser	Ile	Thr	Ile	Leu	Ser	Thr	Pro	Glu	Gly	Thr	Ser	Ala	Ala
245				250				255							
Cys	Lys	Ser	Ile	Leu	Glu	Ile	Met	His	Lys	Glu	Ala	Gln	Asp	Ile	Lys
260				265				270							
Phe	Thr	Glu	Glu	Ile	Pro	Leu	Lys	Ile	Leu	Ala	His	Asn	Asn	Phe	Val
275				280				285							
Gly	Arg	Leu	Ile	Gly	Lys	Glu	Gly	Arg	Asn	Leu	Lys	Lys	Ile	Glu	Gln
290				295				300							
Asp	Thr	Asp	Thr	Lys	Ile	Thr	Ile	Ser	Pro	Leu	Gln	Glu	Leu	Thr	Leu
305				310				315							
Tyr	Asn	Pro	Glu	Arg	Thr	Ile	Thr	Val	Lys	Gly	Asn	Val	Glu	Thr	Cys
325				330				335							
Ala	Lys	Ala	Glu	Glu	Glu	Ile	Met	Lys	Lys	Ile	Arg	Glu	Ser	Tyr	Glu
340				345				350							
Asn	Asp	Ile	Ala	Ser	Met	Asn	Leu	Gln	Ala	His	Leu	Ile	Pro	Gly	Leu
355				360				365							
Asn	Leu	Asn	Ala	Leu	Gly	Leu	Phe	Pro	Pro	Thr	Ser	Gly	Met	Pro	Pro
370				375				380							

156

565 570 575

Arg Arg Lys

<210> 349  
 <211> 207  
 <212> DNA  
 <213> Homo sapiens

<400> 349  
 atgtggcagc cccctcttctt caagtggctc ttgtctctgtt gccctgggag ttctcaaatt 60  
 gctgcagcag cctccaccca gccctaggat gacatcaata cacagaggaa gaagagtcag 120  
 gaaaagatga gagaagtac agactctcct gggcgacccc gagagcttac cattctctcag 180  
 acttctctcac atggtgctaa cagattt 207

<210> 350  
 <211> 69  
 <212> PRT  
 <213> Homo sapiens

<400> 350  
 Met Trp Gln Pro Leu Phe Phe Lys Trp Leu Leu Ser Cys Cys Pro Gly  
 1 5 10 15  
 Ser Ser Gln Ile Ala Ala Ala Ser Thr Gln Pro Glu Asp Asp Ile  
 20 25 30  
 Asn Thr Gln Arg Lys Lys Ser Gln Glu Lys Met Arg Glu Val Thr Asp  
 35 40 45  
 Ser Pro Gly Arg Pro Arg Glu Leu Thr Ile Pro Gln Thr Ser Ser His  
 50 55 60  
 Gly Ala Asn Arg Phe  
 65

<210> 351  
 <211> 1012  
 <212> DNA  
 <213> Homo sapiens

<400> 351  
 cccctctagaa ataattttgt ttaacttttaa gaaggagata tacatatgca tcaccatcac 60  
 catcacacgg cgcgctccga taacttccag ctgtcccagg gtgggcaagg attcgccatt 120  
 ccgatcgggc aggcgatggc gatcgccggc cagatcaacg tccccacgt tcatatcggg 180  
 cctaccggct tctcggcgtt ggggtgtgtc gacaacaacg gcaacggcgc acgagtcaca 240  
 cgcgtgtctg ggcgcgtccg ggcggcaagt ctccgcatct ccaccggcga cgtgatcacc 300  
 gcggtcgacg gcgctccgat caactcggcc accgcgatgg cggacgcgct taacggggcat 360  
 catcccggtg acgtcatctc ggtgacctgg caaaccaagt cggcgccgac gcgtacaggg 420  
 aacgtgacat tggccgaggg acccccggcc gaattcatgg attgggggac gctgcacact 480  
 ttcatcgggg gtgtcaacaa acactccacc agcatcggga aggtgtggat cacagtcate 540  
 tttattttcc gagtcatgat cctcgtgggt gctgccagg aagtgtgggg tgacgagcaa 600  
 gaggactctg tctgcaacac actgcaaccc gcatgcaaaa atgtgtgcta tgaccacttt 660  
 ttcccggtgt cccacatccg gctgtgggccc ctccagctga tcttgtctc caccccagcg 720  
 ctgctgtgtg ccatgcatgt ggccacttac aggcacgaaa ccactcgcaa gttcaggcga 780  
 ggagagaaga ggaatgattt caaagacata gaggacatta aaaagcagaa ggttcggata 840  
 gaggggtgac tcgagcacca ccaccaccac cactgagatc cggctgctaa caaagcccca 900  
 aaggaagctg agttggctgc tgcaccggct gagcaataac tagcataacc ccttgggggc 960  
 tctaaacggg tcttgagggg ttttttgctg aaaggaggaa ctatatccgg at 1012



<210> 352  
 <211> 267  
 <212> PRT  
 <213> Homo sapiens

<400> 352

```

Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
 1          5          10          15
Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
 20
Ile Ala Gly Gln Ile Lys Leu Pro Thr Val His Ile Gly Pro Thr Ala
 35          40          45
Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val
 50          55          60
Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr
 65          70          75
Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr
 85          90          95
Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser
100          105          110
Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr
115          120          125
Leu Ala Glu Gly Pro Pro Ala Glu Phe Met Asp Trp Gly Thr Leu His
130          135          140
Thr Phe Ile Gly Gly Val Asn Lys His Ser Thr Ser Ile Gly Lys Val
145          150          155
Trp Ile Thr Val Ile Phe Ile Phe Arg Val Met Ile Leu Val Val Ala
165          170          175
Ala Gln Glu Val Trp Gly Asp Glu Gln Glu Asp Phe Val Cys Asn Thr
180          185          190
Leu Gln Pro Gly Cys Lys Asn Val Cys Tyr Asp His Phe Phe Pro Val
195          200          205
Ser His Ile Arg Leu Trp Ala Leu Gln Leu Ile Phe Val Ser Thr Pro
210          215          220
Ala Leu Leu Val Ala Met His Val Ala Tyr Tyr Arg His Glu Thr Thr
225          230          235
Arg Lys Phe Arg Arg Gly Glu Lys Arg Asn Asp Phe Lys Asp Ile Glu
245          250          255
Asp Ile Lys Lys Gln Lys Val Arg Ile Glu Gly
260          265

```

<210> 353  
 <211> 900  
 <212> DNA  
 <213> Homo sapiens

<400> 353

```

atgcatacc atcaccatca caccggccgcg tccgataact tccagctgtc ccagggtggg 60
cagggtattc ccatctcgat cgggcaggcg atggcgatcg cgggccagat caagcttccc 120
accgttcata tcgggcctac cgccttcctc ggcttgggtg ttgtcgacaa caacggcaac 180
ggcgcaacga tccaacgcgt ggtcgggagc gctccggcgg caagtctcgg catctccacc 240
ggcgacgtga tcacccgcgt cgacggcgct ccgatcaact cggccaccgc gatggcggac 300
ggccttaacg ggcatcatcc cggtagcgtc atctcgtgtg cctggcaaac caagtccggc 360
ggcacgcgta cagggaacgt gacattggcc gagggacccc cggccgaatt ccacgaaacc 420
actcgcaagt tcaggcgagg agagaagagc aatgatttca aagacataga ggacattaaa 480
aagcagaagg ttccgataga ggggtcgctg tgggtggacgt acaccagcag catctttttc 540
cgaatcatct ttgaagcagc ctttatgtat gtgttttact tcctttacaa tgggtaccac 600

```

```

ctgcctctggg tgttgaaatg tgggattgac cctgccccca accttggtta ctgttttatt 660
tctaggccaa cagagaagac cgtgttttacc atttttatga tttctgcgtc tgtgtattgc 720
atgctgctta acgtggcaga gttgtgctac ctgctgctga aagtgtgttt taggagatca 780
aagagagcac agacgcaaaa aaatcaccoc aatcatgcc taaaggagag taagcagaat 840
gaaatgaatg agctgatttc agatagtggt caaaatgcaa tcacagggtt cccaagctaa 900

```

&lt;210&gt; 354

&lt;211&gt; 299

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 354

```

Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
 1          5          10          15
Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
          20          25          30
Ile Ala Gly Gln Ile Lys Leu Pro Thr Val His Ile Gly Pro Thr Ala
          35          40          45
Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val
          50          55          60
Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr
          65          70          75          80
Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr
          85          90          95
Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser
          100          105          110
Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr
          115          120          125
Leu Ala Glu Gly Pro Pro Ala Glu Phe His Glu Thr Thr Arg Lys Phe
          130          135          140
Arg Arg Gly Glu Lys Arg Asn Asp Phe Lys Asp Ile Glu Asp Ile Lys
          145          150          155          160
Lys Gln Lys Val Arg Ile Glu Gly Ser Leu Trp Trp Thr Tyr Thr Ser
          165          170          175
Ser Ile Phe Phe Arg Ile Ile Phe Glu Ala Ala Phe Met Tyr Val Phe
          180          185          190
Tyr Phe Leu Tyr Asn Gly Tyr His Leu Pro Trp Val Leu Lys Cys Gly
          195          200          205
Ile Asp Pro Cys Pro Asn Leu Val Asp Cys Phe Ile Ser Arg Pro Thr
          210          215          220
Glu Lys Thr Val Phe Thr Ile Phe Met Ile Ser Ala Ser Val Ile Cys
          225          230          235          240
Met Leu Leu Asn Val Ala Glu Leu Cys Tyr Leu Leu Leu Lys Val Cys
          245          250          255
Phe Arg Arg Ser Lys Arg Ala Gln Thr Gln Lys Asn His Pro Asn His
          260          265          270
Ala Leu Lys Glu Ser Lys Gln Asn Glu Met Asn Glu Leu Ile Ser Asp
          275          280          285
Ser Gly Gln Asn Ala Ile Thr Gly Phe Pro Ser
          290          295

```

&lt;210&gt; 355

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 355

ggagtacagc ttcaagacaa tggg

24

&lt;210&gt; 356

&lt;211&gt; 31

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 356

ccatgggaat tcattataat aattttgttc c

31

&lt;210&gt; 357

&lt;211&gt; 920

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 357

Met	Gln	His	His	His	His	His	Gly	Val	Gln	Leu	Gln	Asp	Asn	Gly	
1			5					10					15		
Tyr	Asn	Gly	Leu	Leu	Ile	Ala	Ile	Asn	Pro	Gln	Val	Pro	Glu	Asn	Gln
			20				25						30		
Asn	Leu	Ile	Ser	Asn	Ile	Lys	Glu	Met	Ile	Thr	Glu	Ala	Ser	Phe	Tyr
			35				40					45			
Leu	Phe	Asn	Ala	Thr	Lys	Arg	Arg	Val	Phe	Phe	Arg	Asn	Ile	Lys	Ile
			50			55					60				
Leu	Ile	Pro	Ala	Thr	Trp	Lys	Ala	Asn	Asn	Ser	Lys	Ile	Lys	Gln	
			65			70				75				80	
Glu	Ser	Tyr	Glu	Lys	Ala	Asn	Val	Ile	Val	Thr	Asp	Trp	Tyr	Gly	Ala
			85						90				95		
His	Gly	Asp	Asp	Pro	Tyr	Thr	Leu	Gln	Tyr	Arg	Gly	Cys	Gly	Lys	Glu
			100				105						110		
Gly	Lys	Tyr	Ile	His	Phe	Thr	Pro	Asn	Phe	Leu	Leu	Asn	Asp	Asn	Leu
			115				120					125			
Thr	Ala	Gly	Tyr	Gly	Ser	Arg	Gly	Arg	Val	Phe	Val	His	Glu	Trp	Ala
			130			135				140					
His	Leu	Arg	Trp	Gly	Val	Phe	Asp	Glu	Tyr	Asn	Asn	Asp	Lys	Pro	Phe
			145			150				155					160
Tyr	Ile	Asn	Gly	Gln	Asn	Gln	Ile	Lys	Val	Thr	Arg	Cys	Ser	Ser	Asp
			165						170				175		
Ile	Thr	Gly	Ile	Phe	Val	Cys	Glu	Lys	Gly	Pro	Cys	Pro	Gln	Glu	Asn
			180				185						190		
Cys	Ile	Ile	Ser	Lys	Leu	Phe	Lys	Glu	Gly	Cys	Thr	Phe	Ile	Tyr	Asn
			195				200					205			
Ser	Thr	Gln	Asn	Ala	Thr	Ala	Ser	Ile	Met	Phe	Met	Gln	Ser	Leu	Ser
			210			215				220					
Ser	Val	Val	Glu	Phe	Cys	Asn	Ala	Ser	Thr	His	Asn	Gln	Glu	Ala	Pro
			225			230				235				240	
Asn	Leu	Gln	Asn	Gln	Met	Cys	Ser	Leu	Arg	Ser	Ala	Trp	Asp	Val	Ile
			245						250					255	
Thr	Asp	Ser	Ala	Asp	Phe	His	His	Ser	Phe	Pro	Met	Asn	Gly	Thr	Glu
			260				265					270			
Leu	Pro	Pro	Pro	Pro	Thr	Phe	Ser	Leu	Val	Glu	Ala	Gly	Asp	Lys	Val

275	280	285			
Val Cys Leu Val Leu Asp Val Ser Ser Lys Met Ala Glu Ala Asp Arg					
290	295	300			
Leu Leu Gln Leu Gln Gln Ala Glu Phe Tyr Leu Met Gln Ile Val					
305	310	315			320
Glu Ile His Thr Phe Val Gly Ile Ala Ser Phe Asp Ser Lys Gly Glu					
	325	330			335
Ile Arg Ala Gln Leu His Gln Ile Asn Ser Asn Asp Asp Arg Lys Leu					
	340	345			350
Leu Val Ser Tyr Leu Pro Thr Thr Val Ser Ala Lys Thr Asp Ile Ser					
	355	360			365
Ile Cys Ser Gly Leu Lys Lys Gly Phe Glu Val Val Glu Lys Leu Asn					
	370	375			380
Gly Lys Ala Tyr Gly Ser Val Met Ile Leu Val Thr Ser Gly Asp Asp					
385	390	395			400
Lys Leu Leu Gly Asn Cys Leu Pro Thr Val Leu Ser Ser Gly Ser Thr					
	405	410			415
Ile His Ser Ile Ala Leu Gly Ser Ser Ala Ala Pro Asn Leu Glu Glu					
	420	425			430
Leu Ser Arg Leu Thr Gly Gly Leu Lys Phe Phe Val Pro Asp Ile Ser					
	435	440			445
Asn Ser Asn Ser Met Ile Asp Ala Phe Ser Arg Ile Ser Ser Gly Thr					
	450	455			460
Gly Asp Ile Phe Gln Gln His Ile Gln Leu Glu Ser Thr Gly Glu Asn					
465	470	475			480
Val Lys Pro His His Gln Leu Lys Asn Thr Val Thr Val Asp Asn Thr					
	485	490			495
Val Gly Asn Asp Thr Met Phe Leu Val Thr Trp Gln Ala Ser Gly Pro					
	500	505			510
Pro Glu Ile Ile Leu Phe Asp Pro Asp Gly Arg Lys Tyr Trp Thr Asn					
	515	520			525
Asn Phe Ile Thr Asn Leu Thr Phe Arg Thr Ala Ser Leu Trp Ile Pro					
	530	535			540
Gly Thr Ala Lys Pro Gly His Trp Thr Tyr Thr Leu Asn Asn Thr His					
545	550	555			560
His Ser Leu Gln Ala Leu Lys Val Thr Val Thr Ser Arg Ala Ser Asn					
	565	570			575
Ser Ala Val Pro Pro Ala Thr Val Glu Ala Phe Val Glu Arg Asp Ser					
	580	585			590
Leu His Phe Pro His Pro Val Met Ile Tyr Ala Asn Val Lys Gln Gly					
	595	600			605
Phe Tyr Pro Ile Leu Asn Ala Thr Val Thr Ala Thr Val Glu Pro Glu					
	610	615			620
Thr Gly Asp Pro Val Thr Leu Arg Leu Leu Asp Asp Gly Ala Gly Ala					
625	630	635			640
Asp Val Ile Lys Asn Asp Gly Ile Tyr Ser Arg Tyr Phe Phe Ser Phe					
	645	650			655
Ala Ala Asn Gly Arg Tyr Ser Leu Lys Val His Val Asn His Ser Pro					
	660	665			670
Ser Ile Ser Thr Pro Ala His Ser Ile Pro Gly Ser His Ala Met Tyr					
	675	680			685
Val Pro Gly Tyr Thr Ala Asn Gly Asn Ile Gln Met Asn Ala Pro Arg					
	690	695			700
Lys Ser Val Gly Arg Asn Glu Glu Arg Lys Trp Gly Phe Ser Arg					
705	710	715			720
Val Ser Ser Gly Gly Ser Phe Ser Val Leu Gly Val Pro Ala Gly Pro					
	725	730			735
His Pro Asp Val Phe Pro Pro Cys Lys Ile Ile Asp Leu Glu Ala Val					

Lys	Val	740	Glu	Glu	Leu	Thr	Leu	Ser	Trp	Thr	Ala	Pro	Gly	Glu	Asp
		755						760				765			
Phe	Asp	Gln	Gly	Gln	Ala	Thr	Ser	Tyr	Glu	Ile	Arg	Met	Ser	Lys	Ser
	770					775					780				
Leu	Gln	Asn	Ile	Gln	Asp	Asp	Phe	Asn	Asn	Ala	Ile	Leu	Val	Asn	Thr
785					790					795				800	
Ser	Lys	Arg	Asn	Pro	Gln	Gln	Ala	Gly	Ile	Arg	Glu	Ile	Phe	Thr	Phe
				805					810					815	
Ser	Pro	Gln	Ile	Ser	Thr	Asn	Gly	Pro	Glu	His	Gln	Pro	Asn	Gly	Glu
			820					825					830		
Thr	His	Glu	Ser	His	Arg	Ile	Tyr	Val	Ala	Ile	Arg	Ala	Met	Asp	Arg
		835					840					845			
Asn	Ser	Leu	Gln	Ser	Ala	Val	Ser	Asn	Ile	Ala	Gln	Ala	Pro	Leu	Phe
		850				855						860			
Ile	Pro	Pro	Asn	Ser	Asp	Pro	Val	Pro	Ala	Arg	Asp	Tyr	Leu	Ile	Leu
865					870					875				880	
Lys	Gly	Val	Leu	Thr	Ala	Met	Gly	Leu	Ile	Gly	Ile	Ile	Cys	Leu	Ile
					885				890					895	
Ile	Val	Val	Thr	His	His	Thr	Leu	Ser	Arg	Lys	Lys	Arg	Ala	Asp	Lys
			900					905					910		
Lys	Glu	Asn	Gly	Thr	Lys	Leu	Leu								
		915					920								

```
<210> 358
<211> 2773
<212> DNA
<213> Homo sapiens
```

A400> 358					
catgtcaggc	atcaccacca	tccaccgcga	gtacagcttc	aagacaatg	gtataatgga 60
ttgtgcattg	caattaatcc	tcaggtacct	gagaatcaga	acctcatctc	aaacattaag 120
gaagttagtg	ctgaagttct	atttttacct	tttaattgcta	ccaagagaag	agtatTTTTT 180
agaatatata	cgaggtttaat	acctgccaca	tggaaaagcta	ataataaacg	caaaaaaAa 240
caagaataac	atgaaaaggc	aatgtcata	gtgactgact	gggtatgggc	acatgggat 300
gcacocatca	cotcataata	tcagaggggt	ggaaaagagg	gaaaatacat	tcatttccaca 360
cctaattttc	tactgaatga	taactaatac	gtgtgtcact	gcatacgaag	ccagattgtt 420
gtccatgagt	ggggcccact	ccgttggggt	gttgtctgat	agtataacca	tgacaaaact 480
ttctacataa	atggggcaaaa	tcaaataata	gtgacaaagt	gttctcatctc	catcacaggo 540
atttttgtgt	gtgaaaagg	tctttcccc	caagaaaact	gtattattag	taagcttttt 600
aaagaaggat	gcacctttat	ctcaataatg	accocaaatg	caactgcgat	aaataagtgt 660
atgcgaagtg	tatctctctg	ggttgtaatt	tgtaatgcaa	gtaccoccaa	cgaaagaga 720
coaaactctc	agAACcagat	gtcagccttc	agaagtgcata	gggatgtaat	cacagactct 780
ctgacttctt	accacagctt	tcccatgaac	gggactgagc	ttccacctct	tcccactctt 840
tcgcttgtag	aggctggtga	caaagtggtc	tgtttagtgc	tggatgtgtc	cagcaagatg 900
gcagaggctg	acagactoct	tcaactacaa	caagccgcag	aaattttatt	gatgcagatt 960
ttgtgaattc	atactcttgt	gggcatgtgc	agtttcagca	gaaaggagga	gcatacagct 1020
cagctacacg	aaattaaacg	caatgatgat	cgaaagtgtc	tggtttcaata	tctgccacc 1080
actgtctacg	ctaaaaacag	ctacagcatt	tgttcaggcg	tttaagaaga	atttgaggtg 1140
gttgaaaaac	tgaatggaaa	agcttatggc	tctgtgatga	tattagtgtc	cagcggagat 1200
gataagcttc	ttggcCaatt	atctcccaac	gtgtgcacga	tgattgtcaac	aatctactcc 1260
atttgcctgt	ttctcatctc	agccccaaat	ctggaggaaat	tatcacgtct	tacaggaggt 1320
ttaaagttct	ttgttccaga	tatatccagc	tccaatatga	tgcttgatgc	tttcagttaa 1380
atttctctct	gaactggaga	catTTTCCaC	caacatatgc	agottgaaag	tacaggtgtgaa 1440
aatgtcaaac	ctcaccatca	attgaaaaac	acagtgtact	tggataatac	tgtgggcaac 1500
gacactatgt	tcttagttac	gtgcaggcgc	agttggctct	ctagagattat	attatttgat 1560
ctcgtatggc	gaaataacta	cacaaaataat	tttatcacca	atcgaacttt	tcqacagact 1620

```

agtcttttggga ttccaggaac agctaagcct gggcactgga cttaccacct gaacaatacc 1680
catcattctc tgcaagccct gaaagtgaca gtgacctctc gcgcctccaa ctcagctgtg 1740
ccccagcca ctgtggaagc ctttgtggaa agagacagcc tccattttcc tcatcctgtg 1800
atgatttatg ccaatgtgaa acagggattt tatccattc ttaatgccac tgtcactgcc 1860
acagttgagc cagagactgg agatcctgtt acgctgagac tccttgatga tggagcaggt 1920
gctgatgtta taaaaaatga tggaatttac tcgaggtatt ttttctcctt tgctgcaaat 1980
ggtagataga gcttgaagtc gcatgtcaat cactctccca gcataagcac cccagccacc 2040
tctattccag ggagtcagtc tatgtatgta ccagggttaca cagcaaacgg taatattcag 2100
atgaatgctc caaggaaatc agtaggcaga aatgaggagg agcgaaagtg gggcttttagc 2160
cgagtcagct caggaggtct cttttcagtg ctgggagttc cagctggccc ccacctgat 2220
gtgtttccac catgc aaaat tattgacctg gaagctgtaa aagtagaaga ggaattgacc 2280
ctatcttggga cagcacctgg agaagacttt gatcagggcc aggctacaag ctatgaaata 2340
agaatgagta aaagtctaca gaatatccaa gatgacttta acaatgctat tttagtaaat 2400
acatcaaagc gaaatcctca gcaagctggc atcagggaga tatttacgtt ctacccocaa 2460
atttccagca atggacgtga acatcagcca aatggagaaa cacatgaaag ccacagaatt 2520
tatgttgcaa tacgagcaat ggataggaac tccttacagt ctgctgtatc taacattgcc 2580
caggcgctct tgtttattcc ccccaattct gatcctgtac ctgccagaga ttatcttata 2640
ttgaaaggag ttttaacagc aatgggtttg ataggaatca tttgccttat tatagttgtg 2700
acacatcata ctttaagcag gaaaaagaga gcagacaaga aagagaatgg aacaaaatta 2760
tataatgaa ttc

```

&lt;210&gt; 359

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 359

tggcagcccc tcttcttcaa gtggc

25

&lt;210&gt; 360

&lt;211&gt; 33

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 360

cgccagaatt catcaacaa atctgttagc acc

33

&lt;210&gt; 361

&lt;211&gt; 77

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 361

```

Met Gln His His His His His His Trp Gln Pro Leu Phe Phe Lys Trp
 1          5          10          15
Leu Leu Ser Cys Cys Pro Gly Ser Ser Gln Ile Ala Ala Ala Ala Ser
          20          25          30
Thr Gln Pro Glu Asp Asp Ile Asn Thr Gln Arg Lys Lys Ser Gln Glu
          35          40          45
Lys Met Arg Glu Val Thr Asp Ser Pro Gly Arg Pro Arg Glu Leu Thr
          50          55          60
Ile Pro Gln Thr Ser Ser His Gly Ala Asn Arg Phe Val

```

65

70

75

<210> 362  
 <211> 244  
 <212> DNA  
 <213> Homo sapiens

<400> 362  
 catatgcagc atcaccacca tcaccactgg cagccctct tcttcaagtg gctcttgctc 60  
 tgttgccctg ggagtctca aattgctgca gcagctcca ccagcctga ggatgacatc 120  
 aatacacaga ggaagaagag tcaggaaaag atgagagaag ttacagactc tcctgggcga 180  
 ccccgagagc ttaccattcc tcagacttct tcacatggtg ctaacagatt tgtttgatga 240  
 attc 244

<210> 363  
 <211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 363  
 Met Trp Gln Pro Leu Phe Phe Lys Trp Leu Leu Ser Cys Cys Pro Gly  
 1 5 10 15  
 Ser Ser Gln Ile  
 20

<210> 364  
 <211> 60  
 <212> DNA  
 <213> Homo sapiens

<400> 364  
 atgtggcagc cctcttctt caagtggctc ttgtcctgtt gccctgggag ttctcaaatt 60

<210> 365  
 <211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 365  
 Gly Ser Ser Gln Ile Ala Ala Ala Ala Ser Thr Gln Pro Glu Asp Asp  
 1 5 10 15  
 Ile Asn Thr Gln  
 20

<210> 366  
 <211> 60  
 <212> DNA  
 <213> Homo sapiens

<400> 366  
 gggagttctc aaattgctgc agcagctcc acccagcctg aggatgacat caatacacag 60

<210> 367

<211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 367

Lys Pro Gly His Trp Thr Tyr Thr Leu Asn Asn Thr His His Ser Leu  
 1 5 10 15  
 Gln Ala Leu Lys  
 20

<210> 368  
 <211> 2343  
 <212> DNA  
 <213> Homo sapiens

<400> 368

```

attccggagc gtttgoggct tgcgttcgat gccgtctctc cgcctctcct gggatctgtg 60
gggagctggg gagccgcag cggcccgagg ccggagctgg ccgagccgagc ggagacctgt 120
gcgcgcgcgc tctgaggcgc agcatgtgaa cgggagacgg catccagctgg ggggccgagcc 180
tctcagctgc ccgggattggc taccacggcc gagctcttcg aggagccttt tgtggcagat 240
gaattatatt aacgtcttgt atggagaacc ccaggaggag gctctagagg tggacctgaa 300
gctttttatg ctaaaagatt attagaagaa ttgttaaatc atattcagga actccagata 360
atggatgaaa ggattcagag gaaagtagag aaactagagc aacaatgtca gaaagaagcc 420
aaggaatttg ccaagaaggt acaagagctg cagaaaagca atcaggttgc cttccaacat 480
ttccaagaac tagatgagca cattagctat gtagcaacta aagtctgtca ccttgagagc 540
cagtttagagg gggtaaacac acccagacaa cgggcagctgg aggtcagaaa attgatgaaa 600
tactttaatg agtttctaga tggagaattg aaatctgatg tttttacaaa tcttgaaaag 660
ataaaggaag cagcagacat cattcagaag ttgcacctaa ttgcccaga gttacctttt 720
gatagatatt cagaagttaa atccaaaatt ccaagttaaat accatgattt agaatgccag 780
ctgattcagg agtttaaccag tgctcaaaaga agaggtgaaa totccagaat gagagaagta 840
gcagcagttt tacttcattt taagggttat tccattgtg ttgatgttta tataaagcag 900
tgccaggagg gtgottattt gagaaatgat atatttgaa agctggaaat actctgtcaa 960
agagtgaaca aacaagttgg agatatcttc agtaatccag aaacagctct ggctaaactt 1020
attcaaaatg tatttgaat caaactacag agttttgtga aagagcagttt agaagattct 1080
aggaagtccg atgcagagca atatctcaaa aatctctatg atctgtatcc aagaaccacc 1140
aatctttcgc goaagctgat ggagtttaatt ttaggtaactg ataaacagac ttcttgtct 1200
aagcttatca aatccatttt catttctcat ttggagaact atattgaggt ggagacttga 1260
tatttgaaaa gcagaagtg tatgatccca cagcgtctatt atgattcgaa aaacctcaa 1320
aagagatcca ttggcacagg aggtattcaa aagatttag aaagaaatag acagcgtacc 1380
aacttaacc ttgggccaag tatcgatact catggggaga cttttctatc ccaagaagtg 1440
gtggttaatc ttttaacaaga aacccaaacaa gcctttgaaa gatgtcaatg gctctctgat 1500
cctctgactc taccaggaat tgcttcaga atttttccca ttctgttga atttttatgt 1560
attgagcata ttgattatgc ttggaaaaca ggacttgctg gaattccctc ttcagattct 1620
aggaatgcga atctttattt ttggacggt ttgcaacagg ccaatactat tttcatcttt 1680
tttgacaac agtttaatga tcaactatg ccaactaata gcctctctcc taagtattct 1740
gaatgccttc agaacaagaa agaataatt aaacaaatgg agatgaaatt tggatactgc 1800
attgatagga cattaaatgt tatgatttga cagatgaagc atattttggc tgcagacaag 1860
aagaaaacag attttaagcc agaagatgaa acaactgttt tgattcaata tctactagcc 1920
tgtgtaaaa gctgtgctta cgttaagaaaa caagtggaga agattaaaaa ttccatggat 1980
gggaagaatg tggatacagt ttgatggaa ctggagtag gttttccatg acttatctat 2040
gagcatcttc acaaatatc ctacagttgt ctgggtgcag ttgtggocat tggtagtga 2100
gcgaatata ggaagtgtgc caaagacttc aagattccaa ttgtattaca tctttttgat 2160
actctcatg ctctttgcaa tcttctgtga gttgcccagc ataatttaaa gcaagctctg 2220
tcaggagaac aacttgctaa tctggacaag aatatacttc actcctctgt acaactctgt 2280
gctgattata gatctgccg ccttgctcga caactcagct gagattgaat ttacaaagga 2340
att 2343

```



<210> 369  
 <211> 708  
 <212> PRT  
 <213> Homo sapiens

<400> 369

```

Met Ala Thr Thr Ala Glu Leu Phe Glu Glu Pro Phe Val Ala Asp Glu
 1      5      10      15
Tyr Ile Glu Arg Leu Val Trp Arg Thr Pro Gly Gly Ser Arg Gly
 20      25      30
Gly Pro Glu Ala Phe Asp Pro Lys Arg Leu Leu Glu Glu Phe Val Asn
 35      40      45
His Ile Gln Glu Leu Gln Ile Met Asp Glu Arg Ile Gln Arg Lys Val
 50      55      60
Glu Lys Leu Glu Gln Cys Gln Lys Glu Ala Lys Glu Phe Ala Lys
 65      70      75      80
Lys Val Gln Glu Leu Gln Lys Ser Asn Glu Val Ala Phe Gln His Phe
 85      90      95
Gln Glu Leu Asp Glu His Ile Ser Tyr Val Ala Thr Lys Val Cys His
 100      105      110
Leu Gly Asp Gln Leu Glu Gly Val Asn Thr Pro Arg Gln Arg Ala Val
 115      120      125
Glu Ala Gln Lys Leu Met Lys Tyr Phe Asn Glu Phe Leu Asp Gly Glu
 130      135      140
Leu Lys Ser Asp Val Phe Thr Asn Ser Glu Lys Ile Lys Glu Ala Ala
 145      150      155      160
Asp Ile Ile Gln Lys Leu His Leu Ile Ala Gln Glu Leu Pro Phe Asp
 165      170      175
Arg Phe Ser Glu Val Lys Ser Lys Ile Ala Ser Lys Tyr His Asp Leu
 180      185      190
Glu Cys Gln Leu Ile Gln Glu Phe Thr Ser Ala Gln Arg Arg Gly Glu
 195      200      205
Ile Ser Arg Met Arg Glu Val Ala Ala Val Leu Leu His Phe Lys Gly
 210      215      220
Tyr Ser His Cys Val Asp Val Tyr Ile Lys Gln Cys Gln Glu Gly Ala
 225      230      235      240
Tyr Leu Arg Asn Asp Ile Phe Glu Asp Ala Gly Ile Leu Cys Gln Arg
 245      250      255
Val Asn Lys Gln Val Gly Asp Ile Phe Ser Asn Pro Glu Thr Val Leu
 260      265      270
Ala Lys Leu Ile Gln Asn Val Phe Glu Ile Lys Leu Gln Ser Phe Val
 275      280      285
Lys Glu Gln Leu Glu Glu Cys Arg Lys Ser Asp Ala Glu Gln Tyr Leu
 290      295      300
Lys Asn Leu Tyr Asp Leu Tyr Thr Arg Thr Thr Asn Leu Ser Ser Lys
 305      310      315      320
Leu Met Glu Phe Asn Leu Gly Thr Asp Lys Gln Thr Phe Leu Ser Lys
 325      330      335
Leu Ile Lys Ser Ile Phe Ile Ser Tyr Leu Glu Asn Tyr Ile Glu Val
 340      345      350
Glu Thr Gly Tyr Leu Lys Ser Arg Ser Ala Met Ile Leu Gln Arg Tyr
 355      360      365
Tyr Asp Ser Lys Asn His Gln Lys Arg Ser Ile Gly Thr Gly Gly Ile
 370      375      380
Gln Asp Leu Lys Glu Arg Ile Arg Gln Arg Thr Asn Leu Pro Leu Gly
 385      390      395      400
Pro Ser Ile Asp Thr His Gly Glu Thr Phe Leu Ser Gln Glu Val Val
 405      410      415

```

Val Asn Leu Leu Gln Glu Thr Lys Gln Ala Phe Glu Arg Cys His Arg  
 420 425 430  
 Leu Ser Asp Pro Ser Asp Leu Pro Arg Asn Ala Phe Arg Ile Phe Thr  
 435 440 445  
 Ile Leu Val Glu Phe Leu Cys Ile Glu His Ile Asp Tyr Ala Leu Glu  
 450 455 460  
 Thr Gly Leu Ala Gly Ile Pro Ser Ser Asp Ser Arg Asn Ala Asn Leu  
 465 470 475 480  
 Tyr Phe Leu Asp Val Val Gln Gln Ala Asn Thr Ile Phe His Leu Phe  
 485 490 495  
 Asp Lys Gln Phe Asn Asp His Leu Met Pro Leu Ile Ser Ser Ser Pro  
 500 505 510  
 Lys Leu Ser Glu Cys Leu Gln Lys Lys Glu Ile Ile Glu Gln Met  
 515 520 525  
 Glu Met Lys Leu Asp Thr Gly Ile Asp Arg Thr Leu Asn Cys Met Ile  
 530 535 540  
 Gly Gln Met Lys His Ile Leu Ala Ala Glu Gln Lys Lys Thr Asp Phe  
 545 550 555 560  
 Lys Pro Glu Asp Glu Asn Asn Val Leu Ile Gln Tyr Thr Asn Ala Cys  
 565 570 575  
 Val Lys Val Cys Ala Tyr Val Arg Lys Gln Val Glu Lys Ile Lys Asn  
 580 585 590  
 Ser Met Asp Gly Lys Asn Val Asp Thr Val Leu Met Glu Leu Gly Val  
 595 600 605  
 Arg Phe His Arg Leu Ile Tyr Glu His Leu Gln Gln Tyr Ser Tyr Ser  
 610 615 620  
 Cys Met Gly Gly Met Leu Ala Ile Cys Asp Val Ala Glu Tyr Arg Lys  
 625 630 635 640  
 Cys Ala Lys Asp Phe Lys Ile Pro Met Val Leu His Leu Phe Asp Thr  
 645 650 655  
 Leu His Ala Leu Cys Asn Leu Leu Val Val Ala Pro Asp Asn Leu Lys  
 660 665 670  
 Gln Val Cys Ser Gly Glu Gln Leu Ala Asn Leu Asp Lys Asn Ile Leu  
 675 680 685  
 His Ser Phe Val Gln Leu Arg Ala Asp Tyr Arg Ser Ala Arg Leu Ala  
 690 695 700  
 Arg His Phe Ser  
 705

&lt;210&gt; 370

&lt;211&gt; 60

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 370

gtcaatcact ctcccagcat aagcaccoca gccactcta ttccaggag tcattgctatg 60

&lt;210&gt; 371

&lt;211&gt; 60

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 371

agtagaattt cotctggaac tggagacatt ttccagcaac atattcagct tgaaagtaca 60

<210> 372  
<211> 60  
<212> DNA  
<213> Homo sapiens

<400> 372  
ccagagactg gagatcctgt tacgctgaga ctccttgatg atggagcagg tgctgatgtt 60

<210> 373  
<211> 60  
<212> DNA  
<213> Homo sapiens

<400> 373  
ttacagtctg ctgtatctaa cattgccag ggcctctgt ttattccccc caattctgat 60

<210> 374  
<211> 60  
<212> DNA  
<213> Homo sapiens

<400> 374  
gtgtgcccc cagccactgt ggaagccttt gtggaagag acagcctcca ttttcctcat 60

<210> 375  
<211> 60  
<212> DNA  
<213> Homo sapiens

<400> 375  
aaaaacacag tgactgtgga taatactgtg ggcaacgaca ctatgtttct agttacgtgg 60

<210> 376  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 376  
Leu Gln Ser Ala Val Ser Asn Ile Ala Gln Ala Pro Leu Phe Ile Pro  
1 5 10 15  
Pro Asn Ser Asp  
20

<210> 377  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 377  
Val Asn His Ser Pro Ser Ile Ser Thr Pro Ala His Ser Ile Pro Gly  
1 5 10 15  
Ser His Ala Met  
20

&lt;210&gt; 378

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 378

Pro Glu Thr Gly Asp Pro Val Thr Leu Arg Leu Leu Asp Asp Gly Ala

1 5 10 15

Gly Ala Asp Val

20

&lt;210&gt; 379

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 379

Ala Val Pro Pro Ala Thr Val Glu Ala Phe Val Glu Arg Asp Ser Leu

1 5 10 15

His Phe Pro His

20

&lt;210&gt; 380

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 380

Ser Arg Ile Ser Ser Gly Thr Gly Asp Ile Phe Gln Gln His Ile Gln

1 5 10 15

Leu Glu Ser Thr

20

&lt;210&gt; 381

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 381

Lys Asn Thr Val Thr Val Asp Asn Thr Val Gly Asn Asp Thr Met Phe

1 5 10 15

Leu Val Thr Trp

20

&lt;210&gt; 382

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 382

Lys Pro Gly His Trp Thr Tyr Thr Leu Asn Asn Thr His His Ser Leu

1 5 10 15

Gln Ala Leu Lys  
20

<210> 383  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> PCR primer

<400> 383  
cggcgaattc atggattggg ggacgctgc 29

<210> 384  
<211> 35  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> PCR primer

<400> 384  
cggcctcgag tcacccctct atccgaacct tctgc 35

<210> 385  
<211> 32  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> PCR primer

<400> 385  
cggcgaattc cacgaaccac tcgcaagttc ag 32

<210> 386  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> PCR primer

<400> 386  
cggtctcgagt tagcttgggc ctgtgattgc 30

<210> 387  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 387  
Phe Phe Lys Trp Leu Leu Ser Cys Cys Pro Gly Ser Ser Gln Ile Ala  
1- 5 10 15  
Ala Ala Ala Ser  
20

170

&lt;210&gt; 388

&lt;211&gt; 19

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 388

Leu Ser Cys Cys Pro Gly Ser Ser Gln Ile Ala Ala Ala Ser Thr Gln  
1 5 10 15  
Pro Glu Asp

&lt;210&gt; 389

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 389

Ala Ala Ala Ala Ser Thr Gln Pro Glu Asp Asp Ile Asn Thr Gln Arg  
1 5 10 15  
Lys Lys Ser Gln  
20

&lt;210&gt; 390

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 390

Thr Gln Pro Glu Asp Asp Ile Asn Thr Gln Arg Lys Lys Ser Gln Glu  
1 5 10 15  
Lys Met Arg Glu  
20

&lt;210&gt; 391

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 391

Asp Ile Asn Thr Gln Arg Lys Lys Ser Gln Glu Lys Met Arg Glu Val  
1 5 10 15  
Thr Asp Ser Pro  
20

&lt;210&gt; 392

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 392

Arg Lys Lys Ser Gln Glu Lys Met Arg Glu Val Thr Asp Ser Pro Gly  
1 5 10 15

Arg Pro Arg Glu  
20

<210> 393  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 393  
Glu Lys Met Arg Glu Val Thr Asp Ser Pro Gly Arg Pro Arg Glu Leu  
1 5 10 15  
Thr Ile Pro Gln  
20

<210> 394  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 394  
Val Thr Asp Ser Pro Gly Arg Pro Arg Glu Leu Thr Ile Pro Gln Thr  
1 5 10 15  
Ser Ser His Gly  
20

<210> 395  
<211> 19  
<212> PRT  
<213> Homo sapiens

<400> 395  
Gly Arg Pro Arg Glu Leu Thr Ile Pro Gln Thr Ser Ser His Gly Ala  
1 5 10 15  
Asn Arg Phe

<210> 396  
<211> 19  
<212> PRT  
<213> Homo sapiens

<400> 396  
Met Asn Lys Leu Tyr Ile Gly Asn Leu Ser Glu Asn Ala Ala Pro Ser  
1 5 10 15  
Asp Leu Glu

<210> 397  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 397

172

Ser Glu Asn Ala Ala Pro Ser Asp Leu Glu Ser Ile Phe Lys Asp Ala  
1 5 10 15  
Lys Ile Pro Val  
20

<210> 398  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 398  
Ser Ile Phe Lys Asp Ala Lys Ile Pro Val Ser Gly Pro Phe Leu Val  
1 5 10 15  
Lys Thr Gly Tyr  
20

<210> 399  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 399  
Ser Gly Pro Phe Leu Val Lys Thr Gly Tyr Ala Phe Val Asp Cys Pro  
1 5 10 15  
Asp Glu Ser Trp  
20

<210> 400  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 400  
Ala Phe Val Asp Cys Pro Asp Glu Ser Trp Ala Leu Lys Ala Ile Glu  
1 5 10 15  
Ala Leu Ser Gly  
20

<210> 401  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 401  
Ala Leu Lys Ala Ile Glu Ala Leu Ser Gly Lys Ile Glu Leu His Gly  
1 5 10 15  
Lys Pro Ile Glu  
20

<210> 402  
<211> 20  
<212> PRT  
<213> Homo sapiens



173

&lt;400&gt; 402

Lys Ile Glu Leu His Gly Lys Pro Ile Glu Val Glu His Ser Val Pro  
1 5 10 15  
Lys Arg Gln Arg  
20

&lt;210&gt; 403

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 403

Val Glu His Ser Val Pro Lys Arg Gln Arg Ile Arg Lys Leu Gln Ile  
1 5 10 15  
Arg Asn Ile Pro  
20

&lt;210&gt; 404

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 404

Ile Arg Lys Leu Gln Ile Arg Asn Ile Pro Pro His Leu Gln Trp Glu  
1 5 10 15  
Val Leu Asp Ser  
20

&lt;210&gt; 405

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 405

Ala Val Val Asn Val Thr Tyr Ser Ser Lys Asp Gln Ala Arg Gln Ala  
1 5 10 15  
Leu Asp Lys Leu  
20

&lt;210&gt; 406

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 406

Asp Gln Ala Arg Gln Ala Leu Asp Lys Leu Asn Gly Phe Gln Leu Glu  
1 5 10 15  
Asn Phe Thr Leu  
20

&lt;210&gt; 407

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 407

Asn Gly Phe Gln Leu Glu Asn Phe Thr Leu Lys Val Ala Tyr Ile Pro  
1 5 10 15  
Asp Glu Thr Ala  
20

&lt;210&gt; 408

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 408

Lys Val Ala Tyr Ile Pro Asp Glu Thr Ala Ala Gln Gln Asn Pro Leu  
1 5 10 15  
Gln Gln Pro Arg  
20

&lt;210&gt; 409

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 409

Ala Gln Gln Asn Pro Leu Gln Gln Pro Arg Gly Arg Arg Gly Leu Gly  
1 5 10 15  
Gln Arg Gly Ser  
20

&lt;210&gt; 410

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 410

Gly Arg Arg Gly Leu Gly Gln Arg Gly Ser Ser Arg Gln Gly Ser Pro  
1 5 10 15  
Gly Ser Val Ser  
20

&lt;210&gt; 411

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 411

Ser Arg Gln Gly Ser Pro Gly Ser Val Ser Lys Gln Lys Pro Cys Asp  
1 5 10 15  
Leu Pro Leu Arg  
20

175

<210> 412  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 412  
Lys Gln Lys Pro Cys Asp Leu Pro Leu Arg Leu Leu Val Pro Thr Gln  
1 5 10 15  
Phe Val Gly Ala  
20

<210> 413  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 413  
Leu Leu Val Pro Thr Gln Phe Val Gly Ala Ile Ile Gly Lys Glu Gly  
1 5 10 15  
Ala Thr Ile Arg  
20

<210> 414  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 414  
Ile Ile Gly Lys Glu Gly Ala Thr Ile Arg Asn Ile Thr Lys Gln Thr  
1 5 10 15  
Gln Ser Lys Ile  
20

<210> 415  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 415  
Asn Ile Thr Lys Gln Thr Gln Ser Lys Ile Asp Val His Arg Lys Glu  
1 5 10 15  
Asn Ala Gly Ala  
20

<210> 416  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 416  
Asp Val His Arg Lys Glu Asn Ala Gly Ala Ala Glu Lys Ser Ile Thr  
1 5 10 15  
Ile Leu Ser Thr  
20

<210> 417  
 <211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 417  
 Ala Glu Lys Ser Ile Thr Ile Leu Ser Thr Pro Glu Gly Thr Ser Ala  
 1 5 10 15  
 Ala Cys Lys Ser  
 20

<210> 418  
 <211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 418  
 Pro Glu Gly Thr Ser Ala Ala Cys Lys Ser Ile Leu Glu Ile Met His  
 1 5 10 15  
 Lys Glu Ala Gln  
 20

<210> 419  
 <211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 419  
 Ile Leu Glu Ile Met His Lys Glu Ala Gln Asp Ile Lys Phe Thr Glu  
 1 5 10 15  
 Glu Ile Pro Leu  
 20

<210> 420  
 <211> 455  
 <212> DNA  
 <213> Homo sapiens

<400> 420  
 gaagacatgc ttacttcccc ttcaccttcc ttcattgatg ggggaagagtg ctgcaaccca 60  
 gccctagcca acgccgcatg agaggggagtg tgccgagggc ttctgagaag gtttctctca 120  
 catctagaaa gaagcgctta agatgtggca gccctctctc ttcaagtggc tcttgtcctg 180  
 ttgccctggg agttctcaaa ttgctgcagc agcctccacc cagcctgagg atgacatcaa 240  
 tacacagagg aagaagagtc aggaaaagat gagagaagt acagactctc ctggggcgacc 300  
 cagagagctt accatttctc agactttctc acatgggtgt aacagatttg ttctctaaaag 360  
 taaagctcta gaggcgctca aattggcaat agaagccggg ttccaccata ttgattctgc 420  
 acatgtttac aataatgagg agcaggttgg actgg 455

<210> 421  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 421

actagtgtcc gcgtggcggc ctac

24

&lt;210&gt; 422

&lt;211&gt; 34

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 322

catgagaatt catcacatgc ccttgaagcc tccc

34

&lt;210&gt; 423

&lt;211&gt; 161

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 423

```

Met Gln His His His His His His Thr Ser Val Arg Val Ala Ala
 1          5          10          15
Tyr Phe Glu Asn Phe Leu Ala Ala Trp Arg Pro Val Lys Ala Ser Asp
          20          25          30
Gly Asp Tyr Tyr Thr Leu Ala Val Pro Met Gly Asp Val Pro Met Asp
          35          40          45
Gly Ile Ser Val Ala Asp Ile Gly Ala Ala Val Ser Ser Ile Phe Asn
          50          55          60
Ser Pro Glu Glu Phe Leu Gly Lys Ala Val Gly Leu Ser Ala Glu Ala
          65          70          75          80
Leu Thr Ile Gln Gln Tyr Ala Asp Val Leu Ser Lys Ala Leu Gly Lys
          85          90          95
Glu Val Arg Asp Ala Lys Ile Thr Pro Glu Ala Phe Glu Lys Leu Gly
          100          105          110
Phe Pro Ala Ala Lys Glu Ile Ala Asn Met Cys Arg Phe Tyr Glu Met
          115          120          125
Lys Pro Asp Arg Asp Val Asn Leu Thr His Gln Leu Asn Pro Lys Val
          130          135          140
Lys Ser Phe Ser Gln Phe Ile Ser Glu Asn Gln Gly Ala Phe Lys Gly
          145          150          155          160
Met

```

&lt;210&gt; 424

&lt;211&gt; 489

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 424

```

atgcagcatc accaccatca ccaccacact agtgtccgcg tggcggccta ctttgaaaac 60
tttctcgcgg cgtggcgccc cgtgaaagcc tctgatggag attactacac ctgggtctga 120
ccgatgggag atgtaccaat ggatggtatc tctgttgctg atattggagc agccgtctct 180
agcattttta atttccaga ggaattttta ggcaaggccg tggggctcag tgcagaagca 240
ctaacaatac agcaatatgc tgatgttttg tccaaggcct tggggaaaga agtcagagat 300

```

```

gcaaagatta ccccggaagc ttctgagaag ctgggattcc ctgcagcaaa ggaaatagcc 360
aatatgtgtc gtttctatga aatgaagcca gaccgagatg tcaatctcac ccaccaacta 420
aatcccaaaag tcaaaagctt cagccagttt atctcagaga accaggggagc cttcaagggc 480
atgtgatga                                     489

```

&lt;210&gt; 425

&lt;211&gt; 32

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 425

aacaaaactgt atatcggaaa cctcagcgag aa

32

&lt;210&gt; 426

&lt;211&gt; 33

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 426

ccatagaatt cattaacttcc gtcttgactg agg

33

&lt;210&gt; 427

&lt;211&gt; 586

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 427

```

Met Gln His His His His His Asn Lys Leu Tyr Ile Gly Asn Leu
 1          5          10          15
Ser Glu Asn Ala Ala Pro Ser Asp Leu Glu Ser Ile Phe Lys Asp Ala
          20          25          30
Lys Ile Pro Val Ser Gly Pro Phe Leu Val Lys Thr Gly Tyr Ala Phe
          35          40          45
Val Asp Cys Pro Asp Glu Ser Trp Ala Leu Lys Ala Ile Glu Ala Leu
          50          55          60
Ser Gly Lys Ile Glu Leu His Gly Lys Pro Ile Glu Val Glu His Ser
65          70          75          80
Val Pro Lys Arg Gln Arg Ile Arg Lys Leu Gln Ile Arg Asn Ile Pro
          85          90          95
Pro His Leu Gln Trp Glu Val Leu Asp Ser Leu Leu Val Gln Tyr Gly
          100          105          110
Val Val Glu Ser Cys Glu Gln Val Asn Thr Asp Ser Glu Thr Ala Val
          115          120          125
Val Asn Val Thr Tyr Ser Ser Lys Asp Gln Ala Arg Gln Ala Leu Asp
          130          135          140
Lys Leu Asn Gly Phe Gln Leu Glu Asn Phe Thr Leu Lys Val Ala Tyr
145          150          155          160
Ile Pro Asp Glu Thr Ala Ala Gln Gln Asn Pro Leu Gln Gln Pro Arg
          165          170          175
Gly Arg Arg Gly Leu Gly Gln Arg Gly Ser Ser Arg Gln Gly Ser Pro
          180          185          190
Gly Ser Val Ser Lys Gln Lys Pro Cys Asp Leu Pro Leu Arg Leu Leu

```

```

      195              200              205
Val Pro Thr Gln Phe Val Gly Ala Ile Ile Gly Lys Glu Gly Ala Thr
210              215              220
Ile Arg Asn Ile Thr Lys Gln Thr Gln Ser Lys Ile Asp Val His Arg
225              230              235              240
Lys Glu Asn Ala Gly Ala Ala Glu Lys Ser Ile Thr Ile Leu Ser Thr
245              250              255
Pro Glu Gly Thr Ser Ala Ala Cys Lys Ser Ile Leu Glu Ile Met His
260              265              270
Lys Glu Ala Gln Asp Ile Lys Phe Thr Glu Glu Ile Pro Leu Lys Ile
275              280              285
Leu Ala His Asn Asn Phe Val Gly Arg Leu Ile Gly Lys Glu Gly Arg
290              295              300
Asn Leu Lys Lys Ile Glu Gln Asp Thr Asp Thr Lys Ile Thr Ile Ser
305              310              315              320
Pro Leu Gln Glu Leu Thr Leu Tyr Asn Pro Glu Arg Thr Ile Thr Val
325              330              335
Lys Gly Asn Val Glu Thr Cys Ala Lys Ala Glu Glu Glu Ile Met Lys
340              345              350
Lys Ile Arg Glu Ser Tyr Glu Asn Asp Ile Ala Ser Met Asn Leu Gln
355              360              365
Ala His Leu Ile Pro Gly Leu Asn Leu Asn Ala Leu Gly Leu Phe Pro
370              375              380
Pro Thr Ser Gly Met Pro Pro Pro Thr Ser Gly Pro Pro Ser Ala Met
385              390              395              400
Thr Pro Pro Tyr Pro Gln Phe Glu Gln Ser Glu Thr Glu Thr Val His
405              410              415
Leu Phe Ile Pro Ala Leu Ser Val Gly Ala Ile Ile Gly Lys Gln Gly
420              425              430
Gln His Ile Lys Gln Leu Ser Arg Phe Ala Gly Ala Ser Ile Lys Ile
435              440              445
Ala Pro Ala Glu Ala Pro Asp Ala Lys Val Arg Met Val Ile Ile Thr
450              455              460
Gly Pro Pro Glu Ala Gln Phe Lys Ala Gln Gly Arg Ile Tyr Gly Lys
465              470              475              480
Ile Lys Glu Glu Asn Phe Val Ser Pro Lys Glu Glu Val Lys Leu Glu
485              490              495
Ala His Ile Arg Val Pro Ser Phe Ala Ala Gly Arg Val Ile Gly Lys
500              505              510
Gly Gly Lys Thr Val Asn Glu Leu Gln Asn Leu Ser Ser Ala Glu Val
515              520              525
Val Val Pro Arg Asp Gln Thr Pro Asp Glu Asn Asp Gln Val Val Val
530              535              540
Lys Ile Thr Gly His Phe Tyr Ala Cys Gln Val Ala Gln Arg Lys Ile
545              550              555              560
Gln Glu Ile Leu Thr Gln Val Lys Gln His Gln Gln Gln Lys Ala Leu
565              570              575
Gln Ser Gly Pro Pro Gln Ser Arg Arg Lys
580              585

```

&lt;210&gt; 428

&lt;211&gt; 1764

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 428

atgcagcatc accaccatca ccacaacaaa ctgtatatcg gaaacctcag cgagaacgcc 60

```

gccccctcgg acctagaaga tatcttcaag gacgccaaga tcccggtgtc gggacccttc 120
ctggtagaaga ctggctacgc gttctgtggac tgcccgagac agagctgggc cctcaaggcc 180
atcgaggcgc ttctcagtaa aatagaactg caccgggaac ccataagaagt tgagcactgc 240
gtcccaaaaa ggcaaaagat tcggaaactt cagatacga atatcccgcc tcatttacag 300
tgggaggtgc tggatagttt actagtccag tatggagtgg tggagagctg tgagcaagtg 360
aacactgact cggaactgc agttgtaaat gtaacctatt ccagtaagga ccaagctaga 420
caagcactag acaaaactga tggatttcoag ttagagaatt tcaccttgaa agtagcctat 480
atocctgatg aaacggcgcg ccagcaaaac cccctgcagc agccccgagg tcgcgggggg 540
cttgggcaga ggggctcctc aaggcagggg tctccaggat ccgtatccaa gcagaaacca 600
tgtgatttgc ctctgcgcct gctggttccc acccaatttg ttggagccat cataggaaaa 660
gaaggtgcga ccaatcgga catcacaaa cagaccaggt ctaaaatcga tgtccaccgt 720
aaagaaaatg cgggggctgc tgagaagtcg attactatcc tctctactcc tgaaggcacc 780
tctcgcgctt gtaagtctat tctggagatt atgcataagg aagctcaaga tataaaattc 840
acagaagaga tccccctgaa gattttagct cataataact ttgttgagc tcttattggt 900
aaagaaggaa gaaatcttaa aaaaattgag caagacacag acactaaaat cagcatatct 960
ccattgcagg aattgacgct gtataatcca gaacgacta ttacagttaa agccaatggt 1020
gagacatgtg ccaaagctga ggaggagatc atgaagaaa tcagggagtc ttatgaaaa 1080
gatattgtct ctatgaatct tcaagcacat ttaattcctg gattaaatct gaacgcttg 1140
ggtctgttcc caccacttcc agggatgcca cctccacct caggggcccc ttcagccatg 1200
actctccctc acccgagtt tgagcaatca gaaacggaga gctgttcact gtttatccca 1260
gctctatcag toggtgccat catcggaagc caggggcagc acatcaagca gctttctgcg 1320
tttgctggag ctccaattaa gattgtccca gcggaagcac cagatgtcaa agtgaggatg 1380
gtgattacca ctggacacc agaggctcag ttcaaggctc agggaagaa ttatggaaaa 1440
attaagaag aaaaatttgt tagtctctgga gaagaggtga aacttgaagc tcatatcaga 1500
gtgcatcctc ttgctgctgg cagagttatt ggaagaggag gcaaaacggt gaatgaactt 1560
cagaatttgt caagtgcaga agttgtgtgc agacacctg tgagaaatgac 1620
caagtgtgtg tcaaaaatac tggctcactc tatgcttgc aggtgtccca gagaaaaatt 1680
caggaatttc tgactcaggt aaagcagcac caacaacaga aggcctgtca aagtggacca 1740
cctcagtcga gacggaagta atga 1764

```

&lt;210&gt; 429

&lt;211&gt; 35

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 429.

ccatggaatt cattatttca atataagata atcttc

35

&lt;210&gt; 430

&lt;211&gt; 881

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 430

```

Met Gln His His His His His Gly Val Gln Leu Gln Asp Asn Gly
1          5          10          15
Tyr Asn Gly Leu Leu Ile Ala Ile Asn Pro Gln Val Pro Glu Asn Gln
20          25          30
Asn Leu Ile Ser Asn Ile Lys Glu Met Ile Thr Glu Ala Ser Phe Tyr
35          40          45
Leu Phe Asn Ala Thr Lys Arg Arg Val Phe Phe Arg Asn Ile Lys Ile
50          55          60
Leu Ile Pro Ala Thr Trp Lys Ala Asn Asn Asn Ser Lys Ile Lys Gln
65          70          75          80
Glu Ser Tyr Glu Lys Ala Asn Val Ile Val Thr Asp Trp Tyr Gly Ala

```



His	Gly	Asp	Asp	Pro	Tyr	Thr	Leu	Gln	Tyr	Arg	Gly	Cys	Gly	Lys	Glu
Gly	Lys	Tyr	Ile	His	Phe	Thr	Pro	Asn	Phe	Leu	Leu	Asn	Asp	Asn	Leu
Thr	Ala	Gly	Tyr	Gly	Ser	Arg	Gly	Arg	Val	Phe	Val	His	Glu	Trp	Ala
His	Leu	Arg	Trp	Gly	Val	Phe	Asp	Glu	Tyr	Asn	Asn	Asp	Lys	Pro	Phe
Tyr	Ile	Asn	Gly	Gln	Asn	Gln	Ile	Lys	Val	Thr	Arg	Cys	Ser	Ser	Asp
Ile	Thr	Gly	Ile	Phe	Val	Cys	Glu	Lys	Gly	Pro	Cys	Pro	Gln	Glu	Asn
Cys	Ile	Ile	Ser	Lys	Leu	Phe	Lys	Glu	Gly	Cys	Thr	Phe	Ile	Tyr	Asn
Ser	Thr	Gln	Asn	Ala	Thr	Ala	Ser	Ile	Met	Phe	Met	Gln	Ser	Leu	Ser
Ser	Val	Val	Glu	Phe	Cys	Asn	Ala	Ser	Thr	His	Asn	Gln	Glu	Ala	Pro
Asn	Leu	Gln	Asn	Gln	Met	Cys	Ser	Leu	Arg	Ser	Ala	Trp	Asp	Val	Ile
Thr	Asp	Ser	Ala	Asp	Phe	His	His	Ser	Phe	Pro	Met	Asn	Gly	Thr	Glu
Leu	Pro	Pro	Pro	Pro	Thr	Phe	Ser	Leu	Val	Glu	Ala	Gly	Asp	Lys	Val
Val	Cys	Leu	Val	Leu	Asp	Val	Ser	Ser	Lys	Met	Ala	Glu	Ala	Asp	Arg
Leu	Leu	Gln	Leu	Gln	Gln	Ala	Ala	Glu	Phe	Tyr	Leu	Met	Gln	Ile	Val
Glu	Ile	His	Thr	Phe	Val	Gly	Ile	Ala	Ser	Phe	Asp	Ser	Lys	Gly	Glu
Ile	Arg	Ala	Gln	Leu	His	Gln	Ile	Asn	Ser	Asn	Asp	Asp	Arg	Lys	Leu
Leu	Val	Ser	Tyr	Leu	Pro	Thr	Thr	Val	Ser	Ala	Lys	Thr	Asp	Ile	Ser
Ile	Cys	Ser	Gly	Leu	Lys	Lys	Gly	Phe	Glu	Val	Val	Glu	Lys	Leu	Asn
Gly	Lys	Ala	Tyr	Gly	Ser	Val	Met	Ile	Leu	Val	Thr	Ser	Gly	Asp	Asp
Lys	Leu	Leu	Gly	Asn	Cys	Leu	Pro	Thr	Val	Leu	Ser	Ser	Gly	Ser	Thr
Ile	His	Ser	Ile	Ala	Leu	Gly	Ser	Ser	Ala	Ala	Pro	Asn	Leu	Glu	Glu
Leu	Ser	Arg	Leu	Thr	Gly	Gly	Leu	Lys	Phe	Phe	Val	Pro	Asp	Ile	Ser
Asn	Ser	Asn	Ser	Met	Ile	Asp	Ala	Phe	Ser	Arg	Ile	Ser	Ser	Gly	Thr
Gly	Asp	Ile	Phe	Gln	Gln	His	Ile	Gln	Leu	Glu	Ser	Thr	Gly	Glu	Asn
Val	Lys	Pro	His	Gln	Gln	Leu	Lys	Asn	Thr	Val	Thr	Val	Asp	Asn	Thr
Val	Gly	Asn	Asp	Thr	Met	Phe	Leu	Val	Thr	Trp	Gln	Ala	Ser	Gly	Pro
Pro	Glu	Ile	Ile	Leu	Phe	Asp	Pro	Asp	Gly	Arg	Lys	Tyr	Tyr	Thr	Asn
Asn	Phe	Ile	Thr	Asn	Leu	Thr	Phe	Arg	Thr	Ala	Ser	Leu	Trp	Ile	Pro
Gly	Thr	Ala	Lys	Pro	Gly	His	Trp	Thr	Tyr	Thr	Leu	Asn	Asn	Thr	His

545                      550                      555                      560  
 His Ser Leu Gln Ala Leu Lys Val Thr Val Thr Ser Arg Ala Ser Asn  
                          565                      570                      575  
 Ser Ala Val Pro Ala Thr Val Glu Ala Phe Val Glu Arg Asp Ser  
                          580                      585                      590  
 Leu His Phe Pro His Pro Val Met Ile Tyr Ala Asn Val Lys Gln Gly  
                          595                      600                      605  
 Phe Tyr Pro Ile Leu Asn Ala Thr Val Thr Ala Thr Val Glu Pro Glu  
                          610                      615                      620  
 Thr Gly Asp Pro Val Thr Leu Arg Leu Leu Asp Asp Gly Ala Gly Ala  
 625                      630                      635                      640  
 Asp Val Ile Lys Asn Asp Gly Ile Tyr Ser Arg Tyr Phe Phe Ser Phe  
                          645                      650                      655  
 Ala Ala Asn Gly Arg Tyr Ser Leu Lys Val His Val Asn His Ser Pro  
                          660                      665                      670  
 Ser Ile Ser Thr Pro Ala His Ser Ile Pro Gly Ser His Ala Met Tyr  
                          675                      680                      685  
 Val Pro Gly Tyr Thr Ala Asn Gly Asn Ile Gln Met Asn Ala Pro Arg  
 690                      695                      700  
 Lys Ser Val Gly Arg Asn Glu Glu Glu Arg Lys Trp Gly Phe Ser Arg  
 705                      710                      715                      720  
 Val Ser Ser Gly Gly Ser Phe Ser Val Leu Gly Val Pro Ala Gly Pro  
                          725                      730                      735  
 His Pro Asp Val Phe Pro Pro Cys Lys Ile Ile Asp Leu Glu Ala Val  
                          740                      745                      750  
 Lys Val Glu Glu Glu Leu Thr Leu Ser Trp Thr Ala Pro Gly Glu Asp  
                          755                      760                      765  
 Phe Asp Gln Gly Gln Ala Thr Ser Tyr Glu Ile Arg Met Ser Lys Ser  
                          770                      775                      780  
 Leu Gln Asn Ile Gln Asp Asp Phe Asn Asn Ala Ile Leu Val Asn Thr  
 785                      790                      795                      800  
 Ser Lys Arg Asn Pro Gln Gln Ala Gly Ile Arg Glu Ile Phe Thr Phe  
                          805                      810                      815  
 Ser Pro Gln Ile Ser Thr Asn Gly Pro Glu His Gln Pro Asn Gly Glu  
                          820                      825                      830  
 Thr His Glu Ser His Arg Ile Tyr Val Ala Ile Arg Ala Met Asp Arg  
                          835                      840                      845  
 Asn Ser Leu Gln Ser Ala Val Ser Asn Ile Ala Gln Ala Pro Leu Phe  
 850                      855                      860  
 Ile Pro Pro Asn Ser Asp Pro Val Pro Ala Arg Asp Tyr Leu Ile Leu  
 865                      870                      875                      880  
 Lys

&lt;210&gt; 431

&lt;211&gt; 2646

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 431

atgcagcatc accaccatca ccacggagta cagcttcaag acaatgggta taatggattg 60  
 ctcatctgcaa ttaatcctca ggtacctgag aatcagaacc tcattctcaaa cattaaggaa 120  
 atgataactg aagcttctatt ttacctattt aatgctacca agagaagagt atttttcaga 180  
 aatataaaga ttttaatacc tgccacatgg aaagctaata ataacagcaa aataaaacaa 240  
 gaatcatatg aaaaggcaaa tgtcatagtg actgactggt atgggggcaca tggagatgat 300  
 ccatacaccc tacaatacag aggggtgtga aaagagggaa aatacattca ttcacacct 360  
 aatttctctac tgaatgataa cttaacagct ggctacggat cagcaggccc agtgtttgtc 420

```

catgaatggg cccacotccg ttgggggtgtg ttccgatgagt ataacaatga caaacctttc 480
tacataaatg ggcaaaatca aattaaagt acaagggtgt catctgacat cacaggcatt 540
tttgtgtgtg aaaaagggtcc ttgcccccaa gaaaactgta ttattagtaa gctttttaaa 600
gaaggatgca cotttatcta caatagcacc caaaatgcaa ctgcataaat aatgttcatg 660
caaaagtttat cttctgtgtg tgaattttgt aatgcaagta ccacacaacca agaagcacca 720
aacctacaga accagatgtg cagcctcaga agtgcattgg atgtaatacac agactctgct 780
gactttcoac acagotttcc catgaacggg actgagcttc cactctctcc cactattctgc 840
cttgtagagg ctgggtgaaa agtgggtctgt ttagtgtctg atgtgtccag caagatggca 900
gagggtgaca gactccttca actacaacaa ccgcagaaat tttatttgat gcagattgtt 960
gaaattcata cttctgtggg cattgcacgt ttgcacagca aaggagagat cagagccocag 1020
ctacacaocaa ttaacagcaa tgatgatoga aagttgtctg ttccatatct gccaccact 1080
gtatcagcta aaacagacat cagcatttgt tcagggctta agaaaggatt tgagggtggt 1140
gaaaaactga atggaaaagc ttatggctct gtgatgatat tagtgaccag cggagatgat 1200
aagotttctg gcaattgctt acccaactgt ctcagocagt gtccaacaa tcaactccatt 1260
gcctcgggtt catctgcacg cccaaactct gaggaaatt cactgtttac aggaggttta 1320
aagttcgttt ttccagatat atcaaaactcc aatagcatga ttgatgcttt cagtagaatt 1380
tcctctggaa ctgtgagacat ttccaagcaa catattoacg ttgaaagtac aggtgaaaa 1440
gtcaaacctc accatcaatt gaaaaacaca gtgactgtgg ataactgtg gggcaacgac 1500
actatgtttc tagttacgtg gcaggccagt ggtcctcctg agattatatt atttgatcct 1560
gatggacgaa aatactacac aaataatttt atcaccaatc taacttttcg gacagctagt 1620
ctttggattc caggaaacag taagcctggg cactggactt acaccctgaa caataccact 1680
cattctctgc aagccctgaa agtgacagt accctctcgg cctccaactc agctgtgcoc 1740
ccagccactg tggaaagcct tgtggaaaga gcacagctcc attttctcct tcctgtgatg 1800
atttatgcca atgtgaaaca gggattttat ccatttctta atgccactgt cactgccaca 1860
gttgagccag agactggaga tcctgttaag ctgagactcc ttgatgtgg agcagggtgt 1920
gatgttataa aaaaatgatg aatttactcg aggtattttt tctcctttgc tgcaaaatggt 1980
agatatagtg tgaagtgcaa tgtcaatcac tctccacgca taagcacccc agcccactct 2040
attccaggga gtcatgctat gtatgtacca ggttacacag caaacggtaa tattcagatg 2100
aatgctccaa ggaaatcagt aggcagaaat gaggaggagc gaaagtgggg ctttagcgga 2160
gtcagctcag gaggctcctt ttcagtgtcg ggaagttccag ctggcccoca cctgatgtg 2220
tttccaccat gcaaaattat tgacctggaa gctgtaaaag tagaagagga attgacccta 2280
tottggacag caccctggga agaactttgat caggggccag ctacaagcta tgaataaga 2340
atgagtaaaa gtctacagaa tatccaagat gactttaaca atgtctattt agtaataaca 2400
tcaaagcgaa atcctcagca agctggcctc agggagatat ttacgtttctc accccaaatt 2460
tccacgaatg gacctgaaca tcagccaaat ggagaaacac atgaaagcca cagaatttat 2520
gttgcaatac gagcaatgga taggaaactc ttacagtctg ctgtatctaa cattgcccag 2580
gogcctctgt ttattccccc caattctgat cctgtacctg ccagagatta tcttatattg 2640
aataaa

```

<210> 432  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> PCR primer

<400> 432  
 ogcctgctcg agtcattaat attcatcaga aaatgg

36

<210> 433  
 <211> 371  
 <212> PRT  
 <213> Homo sapiens

<400> 433  
 Met Gln His His His His His Trp Gln Pro Leu Phe Phe Lys Trp  
 1 5 10 15

```

Leu Leu Ser Cys Cys Pro Gly Ser Ser Gln Ile Ala Ala Ala Ala Ser
      20      25      30
Thr Gln Pro Glu Asp Asp Ile Asn Thr Gln Arg Lys Lys Ser Gln Glu
      35      40      45
Lys Met Arg Glu Val Thr Asp Ser Pro Gly Arg Pro Arg Glu Leu Thr
      50      55      60
Ile Pro Gln Thr Ser Ser His Gly Ala Asn Arg Phe Val Pro Lys Ser
      65      70      75      80
Lys Ala Leu Glu Ala Val Lys Leu Ala Ile Glu Ala Gly Phe His His
      85      90      95
Ile Asp Ser Ala His Val Tyr Asn Asn Glu Gln Val Gly Leu Ala
      100      105      110
Ile Arg Ser Lys Ile Ala Asp Gly Ser Val Lys Arg Glu Asp Ile Phe
      115      120      125
Tyr Thr Ser Lys Leu Trp Ser Asn Ser His Arg Pro Glu Leu Val Arg
      130      135      140
Pro Ala Leu Glu Arg Ser Leu Lys Asn Leu Gln Leu Asp Tyr Val Asp
      145      150      155      160
Leu Tyr Leu Ile His Phe Pro Val Ser Val Lys Pro Gly Glu Glu Val
      165      170      175
Ile Pro Lys Asp Glu Asn Gly Lys Ile Leu Phe Asp Thr Val Asp Leu
      180      185      190
Cys Ala Thr Trp Glu Ala Met Glu Lys Cys Lys Asp Ala Gly Leu Ala
      195      200      205
Lys Ser Ile Gly Val Ser Asn Phe Asn His Arg Leu Leu Glu Met Ile
      210      215      220
Leu Asn Lys Pro Gly Leu Lys Tyr Lys Pro Val Cys Asn Gln Val Glu
      225      230      235      240
Cys His Pro Tyr Phe Asn Gln Arg Lys Leu Leu Asp Phe Cys Lys Ser
      245      250      255
Lys Asp Ile Val Leu Val Ala Tyr Ser Ala Leu Gly Ser His Arg Glu
      260      265      270
Glu Pro Trp Val Asp Pro Asn Ser Pro Val Leu Leu Glu Asp Pro Val
      275      280      285
Leu Cys Ala Leu Ala Lys Lys His Lys Arg Thr Pro Ala Leu Ile Ala
      290      295      300
Leu Arg Tyr Gln Leu Gln Arg Gly Val Val Val Leu Ala Lys Ser Tyr
      305      310      315      320
Asn Glu Gln Arg Ile Arg Gln Asn Val Gln Val Phe Glu Phe Gln Leu
      325      330      335
Thr Ser Glu Glu Met Lys Ala Ile Asp Gly Leu Asn Arg Asn Val Arg
      340      345      350
Tyr Leu Thr Leu Asp Ile Phe Ala Gly Pro Pro Asn Tyr Pro Phe Ser
      355      360      365
Asp Glu Tyr
      370

```

&lt;210&gt; 434

&lt;211&gt; 1119

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 434

```

atgcagcatc accaccatca ccaactggcag ccctcttctt tcaagtggct cttgtcctgt      60
tgccctggga gttctcaaat tgcctgcagca gcctccaccc agcctgagga tgacataaat      120
acacagagga agaagagtca ggaagaagtg agagaagtta cagactctcc tgggcgcacc      180
cgagagctta ccatttctca gacttcttca catggtgcta acagatttgt tctctaaaagt      240

```

aaagctctag	aggcogtcaa	attggcaata	gaagccgggt	tccaccatat	tgattctgca	300
catgtttaca	ataatgagga	gcaggttgga	ctggccatcc	gaagcaagat	tcagatggc	360
agtgtaga	gagaagacat	attctacact	tcaaaagcttt	ggagcaattc	ccatcgacca	420
gagttggtcc	gaccagcctt	ggaaaaggtca	ctgaaaaatc	ttcaattgga	ctatgttgac	480
ctctatctta	ttcattttcc	agtgctctgta	aagccaggtg	aggaagtgat	ccaaaaagat	540
gaaaatggaa	aaataactatt	tgacacagtg	gatctctgtg	ccacatggga	ggccatggag	600
aagtgtaaag	atgcaggatt	ggccaagtcc	atcgggggtgt	ccaactctcaa	ccacaggctg	660
ctggagatga	toctcaacaa	gccaggggtc	aagtacaagc	ctgtctgcaa	ccaggtggaa	720
tgcatcctt	acttcaacca	gagaaaactg	ctggatttct	gcaagtcaaa	agacattgtt	780
ctggttgctt	atagtgctct	gggataccat	cgagaagaac	catgggtgga	ccogaactcc	840
ccggtgctct	tgaggagccc	agtcctttgt	gccttggcaa	aaaagcacia	gcgaacccca	900
gccctgattg	ccctgcgcta	ccagctgcag	cgtgggggtg	tggtcctggc	caagagctac	960
aatgagcagc	gcatacgaca	gaacgtgcag	gtgtttgaat	tccagttgac	ttcagaggag	1020
atgaaagcca	tagatggcct	aaacagaaat	tgccgatatt	tgacccttga	tatttttgc	1080
ggccccocta	attatccatt	ttctgatgaa	tattaatga			1119

&lt;210&gt; 435

&lt;211&gt; 36

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Primer

&lt;400&gt; 435

ggatccgccc ccaccatgac atccattoga gctgta 36

&lt;210&gt; 436

&lt;211&gt; 27

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Primer

&lt;400&gt; 436

gtogactcag ctggaccaca gccgcag 27

&lt;210&gt; 437

&lt;211&gt; 37

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Primer

&lt;400&gt; 437

ggatccgccc ccaccatgga ctctggacc ttctgt 37

&lt;210&gt; 438

&lt;211&gt; 27

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Primer

&lt;400&gt; 438

gtcgactcag aaatcctttc tottgac

27

&lt;210&gt; 439

&lt;211&gt; 933

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 439

atggactcct	ggacctctg	ctgtgtgtgc	ctttgcattc	tggtagcaaa	gcacacagat	60
gctggagtta	tccagtcacc	coggcagcag	gtgacagaga	tgggacaaga	agtgactctg	120
agatgtaaac	caatttcagg	acacgactac	ctttcttggt	acagacagac	catgatgcgg	180
ggactggagt	tgctcattta	ctttaacaac	aacgttccga	tagatgattc	agggatgccc	240
gaggatcgat	tctcagctaa	gatgcctaag	gcatacttct	ccactctgaa	gacccagccc	300
tcagaaccca	gggactcagc	tgtgtacttc	tgtgccagca	gtttagtggg	agcaaacact	360
gaagctttct	tggacaagg	caccagactc	acagttgtag	aggacctgaa	caaggtgttc	420
ccaccgcagg	tgctgtgtgt	tgagccatca	gaagcagaga	tctcccacac	ccaaaaggcc	480
caactggtgt	gctcggccac	aggcttcttc	cctgaccacg	tggagctgag	ctgggtgggtg	540
aatgggaagg	aggtgcacag	tggggtcagc	acggaccgcg	agccctcaca	ggagcagccc	600
gcccctaagt	actccagata	ctgcctgagc	agccgcctga	gggtctcgcc	cacctctctg	660
cagaaccccc	gcaaccactt	ccgctgtcaa	gtccagttct	acgggctctc	ggagaatgac	720
gagtggaacc	aggatagggc	caaaccgcgc	accagatctg	tcagcgccga	ggcctggggg	780
agagcagact	gtggctttac	ctcgggtgtc	taccagcaag	gggtcctgtc	tgcccaccatc	840
ctctatgaga	tctcgtcagg	gaaggccacc	ctgtatgctg	tgctgggtcag	cgcccttgtg	900
ttgatggcca	tgttcaagag	aaaggatttc	tga			933

&lt;210&gt; 440

&lt;211&gt; 822

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 440

atgacatcca	ttcagactgt	atttatattc	ctgtggctgc	agctggactt	ggtgaatgga	60
gagaatgtgg	agcagcatcc	ttcaaccctg	agtgtccagg	agggagacag	cgctgttacc	120
aagtgtactt	attcagacag	tgctcacaac	tacttccctt	ggatataagca	agaacttgga	180
aaaagacctc	agcttattat	agacattcgt	tcaaatgtgg	gcgaaaagaa	agaccacaaga	240
attgtctgta	cattgaacaa	gacagccaaa	catttctccc	tgacatcac	agagaccocaa	300
cctgaagact	cggtctgtct	ctctgtgcca	gcaagtatac	tgacacccgg	taaccagttc	360
tattttggga	caggggacaag	tttgacgggtc	attccaaaata	tcacagaccc	tgacctgtgc	420
gtgtaccagg	tgagagactc	taaatccagt	gacaagtctg	tctgctatt	caccgatttt	480
gattctcaaa	caaatgtgtc	acaaagtaag	gattctgatg	tgtatatcac	agacaaaaat	540
gtgtagagca	tgaggtctat	ggacttcaag	agcaacagtg	ctgtggcctg	gagcaacaaa	600
tctgactttg	catgtgcaaa	cgcttcaaac	aacagcatta	ttccagaaga	caccttcttc	660
cccagcccg	aaagttcctg	tgatgtcaag	ctggctcaga	aaagotttga	aacagatacg	720
aacctaaact	ttcaaaacct	gtcagtgatt	gggttcgcga	tcctctcct	gaaagtggcc	780
gggtttaact	tgtcatgac	gctgcggctg	tggtccagct	ga		822

&lt;210&gt; 441

&lt;211&gt; 2311

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 441

gatttaatcc	tatgacaaac	taagttgggt	ctgtcttcac	ctgttttggg	gaggtgtgtg	60
aagagttggg	gtttgtctcag	gaagagattt	aagcatgctt	gcttaccag	actcagagaa	120
gtctccctgt	tctgtcctag	ctatgttctc	gtgtgtgtgt	catttgtatt	ttccagagca	180
aaccgcocag	agtagaagat	ggattggggc	acgtgcaga	cgatcctggg	gggtgtgaac	240
aaacactcca	ccagcattgg	aaagatctgg	ctcacgctcc	tcttcatttt	tcgcattatg	300

```

atcctcgttg  tggctgcaaa  ggaggtgtgg  ggagatgagc  aggccgactt  tgtotgcaac  360
accctgcagc  caggctgcaa  gaacgtgtgc  tacgatcact  acttccccat  ctccccacatc  420
cggtctatgg  cctctgaagt  gatcttcgtg  tccagccagc  cgctcctagt  ggccatgcac  480
gtggcctacc  ggagacatga  gaagaagagg  aagttcatca  agggggagat  aaagagtga  540
tttaaggaca  tcgaggagat  caaaacccag  aaggtccgca  tcgaaggctc  cctgtgtgtg  600
acctacacaa  gcagcatctt  ctccgggtgc  atcttcgaag  ccgccttcac  gtacgtcttc  660
tatgtcatgt  acgacggctt  ctccatgcag  cggtcgtgta  agtgaacagc  ctggccttgt  720
cccaacactg  tggactgtct  tgtgtcccgg  cccacggaga  agactgtctt  cacagtgttc  780
atgattgcag  tgtctggaat  ttgcacctcg  ctgaatgca  ctgaattgtg  ttatttgcga  840
attagatatt  gttctgggaa  gtcaaaaaag  ccagtttaac  gcattgcccc  gttgttagat  900
taagaatatg  acagcatgag  agggatgagg  caaccctgct  tcagctgtca  aggcctagtc  960
gccagcattt  cccaacacaa  agattctgac  cttaaatgca  accatttgaa  acccctgtag  1020
gcctcagggt  aaactccaga  tgccacaatg  agctctgctc  ccctaaagcc  ctcaaaacaaa  1080
ggcctaattc  tatgctgtgc  ttaattttct  ttcaacttaag  ttagtctcac  tgagaccoca  1140
ggctgttagt  ggttatttgt  gtaaggtaact  ttcatatttt  aaacagagga  tatggcatt  1200
tgttctcttc  tctgaggaca  agagaaaaaa  gccaggttcc  acagaggaca  catagaaagt  1260
ttgggtgtcc  tctcggggtt  ctttttgcca  actttcccca  cgttaaaagt  gaacatttgt  1320
tctttcattt  gctttgggaag  ttttaattct  taacagtggg  caaagttacc  agtgctctaa  1380
actctgttac  actttttgga  agtgaaaact  ttgtagtatg  ataggttatt  ttgagttaa  1440
gatgttctgg  ataccattat  atgttcccc  tgtttcagag  gctcagattg  taattgttaa  1500
atggtatgtc  attgcttact  atgatttaat  ttgaaatatg  gtcttttgtt  tatgattact  1560
ttgcagacca  gctgagagag  gctgtctgtt  gtattcattg  tggctcatagc  acctaacac  1620
attgtagcct  caatcgagtg  agacagacta  gaagttccta  gttgcttat  gatgacaaat  1680
ggcctcatgt  caaatattag  atgtaatttt  gtgtaagaaa  tacagactgg  atgtaccacc  1740
aactactacc  tgtaatgaca  ggctgttcca  acacatctcc  cttttccatg  ctgtgtgtgc  1800
cagcatcgga  aagaacgctg  atttaaaag  gtgagcttgg  gaattttatt  gacacagtac  1860
catttaattg  ggagacaaaa  atgggggcca  ggggagggag  aagtttctgt  cgttaaaaa  1920
gagtttgaaa  agactggact  cttaattctg  ttgattaaag  atgagctttt  tctaccttca  1980
aaagtttgtt  tggcttacc  ccttcagcct  ccaatttttt  aagtgaaaat  ataaactaata  2040
acatgtgaaa  agaatagaag  ctaaggttta  gataaatatt  gagcagact  ataggaagat  2100
tgaacctgaa  tattgccatt  atgcttgaca  tggtttccaa  aaaatggtac  tccacatact  2160
tcagtggagg  taagtatttt  cctgttgta  agaatagat  tgtaaaagca  ttttgaata  2220
ataagaata  gctttaatga  tatgcttgta  actaaaataa  ttttgtaatg  tatcaaatatc  2280
atttaaaaca  ttaaaatata  atctctataa  t  2311

```

&lt;210&gt; 442

&lt;211&gt; 226

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 442

```

Met Asp Trp Gly Thr Leu Gln Thr Ile Leu Gly Gly Val Asn Lys His
                5                        10                        15

```

```

Ser Thr Ser Ile Gly Lys Ile Trp Leu Thr Val Leu Phe Ile Phe Arg
                20                        25                        30

```

```

Ile Met Ile Leu Val Val Ala Ala Lys Glu Val Trp Gly Asp Glu Gln
                35                        40                        45

```

```

Ala Asp Phe Val Cys Asn Thr Leu Gln Pro Gly Cys Lys Asn Val Cys
                50                        55                        60

```

```

Tyr Asp His Tyr Phe Pro Ile Ser His Ile Arg Leu Trp Ala Leu Gln
                65                        70                        75                        80

```

```

Leu Ile Phe Val Ser Ser Pro Ala Leu Leu Val Ala Met His Val Ala
                85                        90                        95

```

188

Tyr Arg Arg His Glu Lys Lys Arg Lys Phe Ile Lys Gly Glu Ile Lys  
                   100                                  105                                  110  
 Ser Glu Phe Lys Asp Ile Glu Glu Ile Lys Thr Gln Lys Val Arg Ile  
                   115                                  120                                  125  
 Glu Gly Ser Leu Trp Trp Thr Tyr Thr Ser Ser Ile Phe Phe Arg Val  
                   130                                  135                                  140  
 Ile Phe Glu Ala Ala Phe Met Tyr Val Phe Tyr Val Met Tyr Asp Gly  
                   145                                  150                                  155                                  160  
 Phe Ser Met Gln Arg Leu Val Lys Cys Asn Ala Trp Pro Cys Pro Asn  
                   165                                  170                                  175  
 Thr Val Asp Cys Phe Val Ser Arg Pro Thr Glu Lys Thr Val Phe Thr  
                   180                                  185                                  190  
 Val Phe Met Ile Ala Val Ser Gly Ile Cys Ile Leu Leu Asn Val Thr  
                   195                                  200                                  205  
 Glu Leu Cys Tyr Leu Leu Ile Arg Tyr Cys Ser Gly Lys Ser Lys Lys  
                   210                                  215                                  220  
 Pro Val  
 225

<210> 443  
 <211> 23  
 <212> PRT  
 <213> Homo sapiens

<400> 443  
 Val Lys Leu Cys Gly Ile Asp Pro Cys Pro Asn Leu Val Asp Cys Phe  
                                   5                                  10                                  15

Ile Ser Arg Pro Gly Cys Gly  
                   20

<210> 444  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> PCR primer

<400> 444  
 caatcaggca tgcacaacaa actgtatatc ggaaac

36

<210> 445  
 <211> 30  
 <212> DNA



&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 445

cgtcaagatc ttcattactt cegtcttgac

30

&lt;210&gt; 446

&lt;211&gt; 579

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 446

```

Met Asn Lys Leu Tyr Ile Gly Asn Leu Ser Glu Asn Ala Ala Pro Ser
      5              10              15

Asp Leu Glu Ser Ile Phe Lys Asp Ala Lys Ile Pro Val Ser Gly Pro
      20              25              30

Phe Leu Val Lys Thr Gly Tyr Ala Phe Val Asp Cys Pro Asp Glu Ser
      35              40              45

Trp Ala Leu Lys Ala Ile Glu Ala Leu Ser Gly Lys Ile Glu Leu His
      50              55              60

Gly Lys Pro Ile Glu Val Glu His Ser Val Pro Lys Arg Gln Arg Ile
      65              70              75              80

Arg Lys Leu Gln Ile Arg Asn Ile Pro Pro His Leu Gln Trp Glu Val
      85              90              95

Leu Asp Ser Leu Leu Val Gln Tyr Gly Val Val Glu Ser Cys Glu Gln
      100             105             110

Val Asn Thr Asp Ser Glu Thr Ala Val Val Asn Val Thr Tyr Ser Ser
      115             120             125

Lys Asp Gln Ala Arg Gln Ala Leu Asp Lys Leu Asn Gly Phe Gln Leu
      130             135             140

Glu Asn Phe Thr Leu Lys Val Ala Tyr Ile Pro Asp Glu Thr Ala Ala
      145             150             155             160

Gln Gln Asn Pro Leu Gln Gln Pro Arg Gly Arg Arg Gly Leu Gly Gln
      165             170             175

Arg Gly Ser Ser Arg Gln Gly Ser Pro Gly Ser Val Ser Lys Gln Lys
      180             185             190

Pro Cys Asp Leu Pro Leu Arg Leu Leu Val Pro Thr Gln Phe Val Gly
      195             200             205

Ala Ile Ile Gly Lys Glu Gly Ala Thr Ile Arg Asn Ile Thr Lys Gln
      210             215             220

Thr Gln Ser Lys Ile Asp Val His Arg Lys Glu Asn Ala Gly Ala Ala
      225             230             235             240

```

Glu Lys Ser Ile Thr Ile Leu Ser Thr Pro Glu Gly Thr Ser Ala Ala  
 245 250 255  
 Cys Lys Ser Ile Leu Glu Ile Met His Lys Glu Ala Gln Asp Ile Lys  
 260 265 270  
 Phe Thr Glu Glu Ile Pro Leu Lys Ile Leu Ala His Asn Asn Phe Val  
 275 280 285  
 Gly Arg Leu Ile Gly Lys Glu Gly Arg Asn Leu Lys Lys Ile Glu Gln  
 290 295 300  
 Asp Thr Asp Thr Lys Ile Thr Ile Ser Pro Leu Gln Glu Leu Thr Leu  
 305 310 315 320  
 Tyr Asn Pro Glu Arg Thr Ile Thr Val Lys Gly Asn Val Glu Thr Cys  
 325 330 335  
 Ala Lys Ala Glu Glu Glu Ile Met Lys Lys Ile Arg Glu Ser Tyr Glu  
 340 345 350  
 Asn Asp Ile Ala Ser Met Asn Leu Gln Ala His Leu Ile Pro Gly Leu  
 355 360 365  
 Asn Leu Asn Ala Leu Gly Leu Phe Pro Pro Thr Ser Gly Met Pro Pro  
 370 375 380  
 Pro Thr Ser Gly Pro Pro Ser Ala Met Thr Pro Pro Tyr Pro Gln Phe  
 385 390 395 400  
 Glu Gln Ser Glu Thr Glu Thr Val His Leu Phe Ile Pro Ala Leu Ser  
 405 410 415  
 Val Gly Ala Ile Ile Gly Lys Gln Gly Gln His Ile Lys Gln Leu Ser  
 420 425 430  
 Arg Phe Ala Gly Ala Ser Ile Lys Ile Ala Pro Ala Glu Ala Pro Asp  
 435 440 445  
 Ala Lys Val Arg Met Val Ile Ile Thr Gly Pro Pro Glu Ala Gln Phe  
 450 455 460  
 Lys Ala Gln Gly Arg Ile Tyr Gly Lys Ile Lys Glu Glu Asn Phe Val  
 465 470 475 480  
 Ser Pro Lys Glu Glu Val Lys Leu Glu Ala His Ile Arg Val Pro Ser  
 485 490 495  
 Phe Ala Ala Gly Arg Val Ile Gly Lys Gly Gly Lys Thr Val Asn Glu  
 500 505 510  
 Leu Gln Asn Leu Ser Ser Ala Glu Val Val Val Pro Arg Asp Gln Thr  
 515 520 525  
 Pro Asp Glu Asn Asp Gln Val Val Val Lys Ile Thr Gly His Phe Tyr  
 530 535 540  
 Ala Cys Gln Val Ala Gln Arg Lys Ile Gln Glu Ile Leu Thr Gln Val

545		550		555		560
Lys	Gln	Gln	Gln	Lys	Ala	Leu
		565		570		575
Gln	Ser	Gly	Pro	Pro	Gln	Ser

Arg Arg Lys

&lt;210&gt; 447

&lt;211&gt; 1743

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 447

```

atgaacaaac tgtatatcgg aaacctcagc gagaacgcgc cccctcgcga cctagaaagt 60
atcttcaagg acgccaagat cccgggtgct ggacccttcc tgggtgaagac tggctacgcg 120
ttcgtggact gccggagcga gagctgggcc ctcaaggcca tcgaggcgct ttcagggtaaa 180
atagaaactgc acgggaaacc catagaagtt gagcactcgg tcccaaaaag gcaaaggatt 240
cggaactctc agatacgaag tatcccgctc catttacagt gggagggtgct ggaatagttta 300
ctagtcagat atggagtggt ggagagctgt gagcaagtga acactgactc ggaaactgca 360
gtgttaaatg taacctattc cagtaaggac caagctagac aagcactaga caaactgaat 420
ggatttcagt tagagaattt caccttgaaa gtacgctata tccctgatga aacggccgcc 480
cagcaaaaacc ccttgcagca gccccgaggt cgcggggggc ttggggcagag gggctcctca 540
aggcagggggt cccagagatc cgtatccaag cagaaacctc gtgatttgcc tctgcgcctg 600
ctggttcccc cccaatttgt tggagccatc ataggaaaag aagggtgccc cattcggaac 660
atcacccaaac agaccagtc taaaatcgat gtcccacgta aagaaaatgc gggggctgct 720
gagaagtoga ttactatcct ctctactcct gaaggcaact ctgcgggttg taagtctatt 780
ctggagatta tgcataagga agctcaagat ataaaattca cagaagagat ccccttgaag 840
attttagctc ataataactt tgttggacgt cttatttgta aagaaggaag aaatctttaa 900
aaaattgagc aagacacaga caactaaatc acgatatctc cattgcagga attgacgctg 960
tataatccag aacgcactat tacagttaaa ggcaatgttg agacatgtgc caaagctcag 1020
gaggagatca tgaagaaaat caggggagtct tatgaaaatg atattgcttc tatgaattct 1080
caagcacatt taattccttg attaaatctg aacgccttgg gtctgttccc acccacttca 1140
gggatgccac ctcccacctc agggccctct tcagccatga ctctccctca ccgcagttt 1200
gagcaaatcag aaacggagac tgttcatctg tttatcccag ctctatcagt cggtgccatc 1260
atcggcgaagc agggccagca catcaagcag cttctcgtct ttgctggagc tccaattaag 1320
attgtccagc cggaagcacc agatgctaaa gtgaggatgg tgattatcac tggaccacca 1380
gaggctcagt tcaaggctca ggaagaattt tatggaaaaa ttaagagaag aaactttggt 1440
agtctaaaag aagaggtgaa acttgaagct catatcagag tgcctatcctt tgctgtcggc 1500
agagttattg gaaaaggagg caaaacgggt aatgaacttc agaattttgtc aagtgcagaa 1560
gttgttgtcc ctctgacca gcacactgat gagaatgacc aagtgtgtgt caaaaatac 1620
ggtcactctt atgottgcca ggttgcccag agaaaaattc aggaattctt gactcaggta 1680
aagcagcacc aacaacagaa ggctctgcac atgtggaccac ctcagtcaag acggaagtaa 1740
tga

```

&lt;210&gt; 448

&lt;211&gt; 35

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 448

cgtactagca tatgaacaaa ctgtatatcg gaaac



275		280		285
Gly Arg Leu Ile Gly Lys Glu Gly Arg Asn Leu Lys Lys Ile Glu Gln				
290		295		300
Asp Thr Asp Thr Lys Ile Thr Ile Ser Pro Leu Gln Glu Leu Thr Leu				
305		310		320
Tyr Asn Pro Glu Arg Thr Ile Thr Val Lys Gly Asn Val Glu Thr Cys				
		325		335
Ala Lys Ala Glu Glu Glu Ile Met Lys Lys Ile Arg Glu Ser Tyr Glu				
		340		350
Asn Asp Ile Ala Ser Met Asn Leu Gln Ala His Leu Ile Pro Gly Leu				
		355		365
Asn Leu Asn Ala Leu Gly Leu Phe Pro Pro Thr Ser Gly Met Pro Pro				
		370		380
Pro Thr Ser Gly Pro Pro Ser Ala Met Thr Pro Pro Tyr Pro Gln Phe				
		385		400
Glu Gln Ser Glu Thr Glu Thr Val His Leu Phe Ile Pro Ala Leu Ser				
		405		415
Val Gly Ala Ile Ile Gly Lys Gln Gly Gln His Ile Lys Gln Leu Ser				
		420		430
Arg Phe Ala Gly Ala Ser Ile Lys Ile Ala Pro Ala Glu Ala Pro Asp				
		435		445
Ala Lys Val Arg Met Val Ile Ile Thr Gly Pro Pro Glu Ala Gln Phe				
		450		460
Lys Ala Gln Gly Arg Ile Tyr Gly Lys Ile Lys Glu Glu Asn Phe Val				
		465		480
Ser Pro Lys Glu Glu Val Lys Leu Glu Ala His Ile Arg Val Pro Ser				
		485		495
Phe Ala Ala Gly Arg Val Ile Gly Lys Gly Gly Lys Thr Val Asn Glu				
		500		510
Leu Gln Asn Leu Ser Ser Ala Glu Val Val Val Pro Arg Asp Gln Thr				
		515		525
Pro Asp Glu Asn Asp Gln Val Val Val Lys Ile Thr Gly His Phe Tyr				
		530		540
Ala Cys Gln Val Ala Gln Arg Lys Ile Gln Glu Ile Leu Thr Gln Val				
		545		560
Lys Gln His Gln Gln Gln Lys Ala Leu Gln Ser Gly Pro Pro Gln Ser				
		565		575
Arg Arg Lys				

<210> 450  
 <211> 1743  
 <212> DNA  
 <213> Homo sapiens

<400> 450  
 atgaacaaac tgtatatogg aaacotcago gagaacgocg cccctcgcga cctagaaagt 60  
 atottcaagg acgccaagat cccgggtgoc ggacccttcc tgggtgaagac tggctacgoc 120  
 ttogtggact gcccggaaga gagctgggoc ctcaaggcca tcgaggcgct ttcagggtaaa 180  
 atagaactgc accgggaacc catagaagtt gagcactcgg tcccaaaaag gcaaggatt 240  
 cggaaacttc agatacgaat tatcccgctt catttacagt gggagggtgt ggaatgttta 300  
 ctagtccagt atggagtggt ggagagctgt gagcaagtga acactgactc ggaactgca 360  
 gttgtaaaat taacctatc cagtaaggac caagctagac aagcactaga caaactgaat 420  
 ggatttcagt tagagaattt cacottgaaa gtacgtata tccctgatga aacggccgcc 480  
 cagcaaaacc ccttgcagca gcccgaggt cgcggggggc ttgggcagag gggctcctca 540  
 aggcagggtt ctccaggatc cgtatccaag cagaaacat gtgatttgcc tctgcgctg 600  
 ctgggtccca cccaatttgt tggagccatc ataggaagaag aaggtgccac cattcggaac 660  
 atcaccaaac agacccagtc taaaatcgat gtccaccgta aagaaaatgc ggggctgct 720  
 gagaagtcca ttactatcct ctctactcct gaaggcaact ctgcggtgtg taagtctatt 780  
 ctggagatta tgcataagga agctcaagat ataaaattca cagaagagat ccccttgaag 840  
 attttagctc ataataactt tgttggagct ctatttggtt aagaaggagaa aaactctaaa 900  
 aaaattgagc aagacacaga cactaaaatc acgatatctc cacttgcagga attgacgctg 960  
 tataatccag aacgcactat tacagttaaa ggcaatgttg agacatgtgc caaagctgag 1020  
 caggagatca tgaagaaaat caggaggtct tatgaaaatg atattgcttc tatgaatctt 1080  
 caagcacatt taattcctgg attaaatctg aacgccttgg gtctgttccc acccacttca 1140  
 gggatgccac ctcccacctc agggccccc tccagccatga ctccctccta ccgagagttt 1200  
 gagcaatcag aaacggagac tgttcatctg ttatcccatc ctctatcagt cgggtgccatc 1260  
 atcggcaagc agggccagca catcaagcag ctttctcgtc ttgctggagc ttcaattaa 1320  
 attgctccag cgggaagcacc agatgctaaa gtgaggatgg tgattatcac tggaccacca 1380  
 gaggtcagct tcaaggctca gggaagaatt tatggaaaaa ttaaagaaga aaactttgtt 1440  
 agtccataag aagaggtgaa acttgaagct catatcagag tgccatcctt tgctgctggc 1500  
 agagttattg gaaaaggagg caaaacgggt aatgaacttc agaatttgtc aagtgcagaa 1560  
 gttgttgtcc ctctgacca gacacctgat gagaatgacc aagtgttgtt caaataact 1620  
 ggtcacttct atgcttcca ggttgccagc agaaaaatc aggaaattct gactcaggt 1680  
 aagcagcacc aacaacagaa ggctctgcaa atgggaccac ctcagtcagc acggaagtaa 1740  
 tga 1743

<210> 451  
 <211> 25  
 <212> PRT  
 <213> Homo sapiens

<400> 451  
 Leu Gly Lys Glu Val Arg Asp Ala Lys Ile Thr Pro Glu Ala Phe Glu  
 5 10 15  
 Lys Leu Gly Phe Pro Ala Ala Lys Glu  
 20 25

<210> 452  
 <211> 25  
 <212> PRT  
 <213> Homo sapiens

<400> 452

195

Lys Ala Ser Asp Gly Asp Tyr Tyr Thr Leu Ala Val Pro Met Gly Asp  
5 10 15

Val Pro Met Asp Gly Ile Ser Val Ala  
20 25

<210> 453  
<211> 16  
<212> PRT  
<213> Homo sapiens

<400> 453  
Pro Asp Arg Asp Val Asn Leu Thr His Gln Leu Asn Pro Lys Val Lys  
5 10 15

<210> 454  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 454  
Lys Ile Ala Pro Ala Glu Ala Pro Asp Ala Lys Val Arg Met Val Ile  
5 10 15

Ile Thr Gly Pro  
20

<210> 455  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 455  
Pro Asp Glu Thr Ala Ala Gln Gln Asn Pro Leu Gln Gln Pro Arg Gly  
5 10 15

Arg Arg Gly Leu  
20

<210> 456  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 456  
Arg Thr Ile Thr Val Lys Gly Asn Val Glu Thr Cys Ala Lys Ala Glu  
5 10 15

Glu Glu Ile Met  
20

<210> 457  
<211> 20  
<212> PRT

<213> Homo sapiens

<400> 457

Ala Phe Val Asp Cys Pro Asp Glu Ser Trp Ala Leu Lys Ala Ile Glu  
5 10 15

Ala Leu Ser Gly  
20

<210> 458

<211> 20

<212> PRT

<213> Homo sapiens

<400> 458

Ile Arg Lys Leu Gln Ile Arg Asn Ile Pro Pro His Leu Gln Trp Glu  
5 10 15

Val Leu Asp Ser  
20

<210> 459

<211> 20

<212> PRT

<213> Homo sapiens

<400> 459

Ala Gln Gln Asn Pro Leu Gln Gln Pro Arg Gly Arg Arg Gly Leu Gly  
5 10 15

Gln Arg Gly Ser  
20

<210> 460

<211> 20

<212> PRT

<213> Homo sapiens

<400> 460

Asp Val His Arg Lys Glu Asn Ala Gly Ala Ala Glu Lys Ser Ile Thr  
5 10 15

Ile Leu Ser Thr  
20

<210> 461

<211> 20

<212> PRT

<213> Homo sapiens

<400> 461

Leu Tyr Asn Pro Glu Arg Thr Ile Thr Val Lys Gly Asn Val Glu Thr  
5 10 15



Cys Ala Lys Ala  
20

<210> 462  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 462  
Glu Glu Glu Ile Met Lys Lys Ile Arg Glu Ser Tyr Glu Asn Asp Ile  
5 10 15

Ala Ser Met Asn  
20

<210> 463  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 463  
Leu Asn Ala Leu Gly Leu Phe Pro Pro Thr Ser Gly Met Pro Pro Pro  
5 10 15

Thr Ser Gly Pro  
20

<210> 464  
<211> 20  
<212> PRT  
<213> Homo sapiens

<400> 464  
Lys Ile Ala Pro Ala Glu Ala Pro Asp Ala Lys Val Arg Met Val Ile  
5 10 15

Ile Thr Gly Pro  
20

<210> 465  
<211> 18  
<212> PRT  
<213> Homo sapiens

<400> 465  
Thr Gly Tyr Ala Phe Val Asp Cys Pro Asp Glu Ser Trp Ala Leu Lys Ile  
5 10 15  
Glu

<210> 466  
<211> 11  
<212> PRT  
<213> Homo sapiens

&lt;400&gt; 466

Phe Val Asp Cys Pro Asp Glu Ser Trp Ala Leu  
5 10

&lt;210&gt; 467

&lt;211&gt; 33

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 467

ttcgtggact gcccgacga gagctgggcc etc

33